1 TTAATTAAGGAGGACCATC ATG AAC GAG GCC ATC GCC GTC GGC ATG TCC TGC CGC CTG CCG MNEAIAVV G M 65 AAG GCC TCG AAC CCG GCC GCC TTC TCG GAG CTG CTG CGG AAC GCG GAG AGC GCC GTC ACC ASNPAAFWELLRNGES 16 K 125 GAC GTG CCC TCC GGC CGG TGG ACG TCG GTG CTC GGG GGA GCG GAC GCC GAG GAG CCG GCG 36 D V P G R W T S v L G G A A 185 GAG TCC GGT GTC CGC CGG GGC GGC TTC CTC GAC TCC CTC GAC CTC TTC GAC GCG GCC TTC 244 R G G D 75 245 TTC GGA ATC TCG CCC CGT GAG GCC GCC ATG GAC CCG CAG CAG CGA CTG GTC CTC GAA 304 PREA A A M D P 0 R 305 CTC GCC TGG GAG GCG CTG GAG GAC GCC GGA ATC GTC CCC GGC ACC CTC GCC GGA AGC CGC 364 96 L A W E A L E D A G G А 365 ACC GCC GTC TTC GTC GGC ACC CTG CGG GAC GAC TAC ACG AGC CTC CTC TAC CAG CAC GGC 424 R D D Y T S L L 135 425 GAG CAG GCC ATC ACC CAG CAC ACC ATG GCG GGC GTG AAC CGG GGC GTC ATC GCC AAC CGC 484 136 E Q A о н T M A G N R G 485 GTC TCG TAC CAC CTC GGC CTG CAG GGC CCG AGC CTC ACC GTC GAC GCC GCG CAG TCG TCC 544 G S Т 175 545 TOG CTC GTC GCC GTG CAC CTG GCC TGC GAG TCC CTG CGC GCC GGG GAG TCC ACG ACG GCG 604 HLA C ESLRA 195 605 CTC GTC GCC GGC GTG AAC CTC AAC ATC CTC GCG GAG AGC GCC GTG ACG GAG GAG CGC TTC 664 NLNI L A E S 665 GGT GGA CTC TCC CCG GAC GGC ACC GCC TAC ACC TTC GAC GCG CCG GCC AAC GGA TTC GTC 724 D G T AYTF DARA N 235 725 CGG GGC GAG GGC GGA GTC GTC GTA CTC AAG CCG CTC TCC CGC GCC CTC GCC GAC GGC 784 236 R G E G L K R A 785 GAC CGT GTC CAC GGC GTC ATC CGC GCC AGC GCC GTC AAC AAC GAC GGA GCC ACC CCG GGT 844 I R A A N N D 275 845 CTC ACC GTG CCC AGC AGG GCC GCC CAG GAG AAG GTG CTG CGC GAG GCG TAC CGG AAG GCG 904 RAAQEK V REA 905 GCC CTG GAC CCG TCC GCC GTC CAG TAC GTC GAA CTC CAC GGC ACC GGA ACC CCC GTC GGC 964 H G T G 1024 A A T. C A VLGSARP A D 1025 CCC CTG CTC GTC GGC TCG GCC AAG ACG AAC GTC GGG CAC CTC GAA GGC GCC GCC GGC ATC 1084 G S A KTN v G H L EGAA 355 1085 GTC GGC CTC ATC AAG ACG CTC CTC GGC CTC GGC CGC CGC ATC CCG GCG AGC CTC AAC LLA G 1145 TTC OGT ACG CCC CAC CCG GAC ATC CCG CTC GAC ACC CTC GGG CTC GAC GTG CCC GAC GGC 1204 P L D T L G D 395 1205 CTG CGG GAG TGG CCG CAC CCG GAC CGC GAA CTC CTC GCC GGC GTC AGC TCG TTC GGC ATG 1264 P н Р RELLA D 1324 T N S E G P A 435 1325 GGC ATC GAT GAG GAG ACC CCC GTC GAC AGC GGG GCC GCA CTG CCC TTC GTC ACC GGC E E T D S G A A 1385 CGC GGC GGC GAG GCC CTG CGC GCC CAG GCC CTG CAC GAG GCC GTC GAA GCG GAC 1444 456 R G G E A L R A Q ARR H

SEP 2 4 2004

FIG. 23A

1445 CCG GAG CTC GCG CCC GCC GCA CTC GCC CGG TCG CTG GTC ACC CGT ACG GTC TTC ACG 1504 AALARSL 1505 CAC CGG TOG GTC GTC CTC GCC CCG GAC CGC CGC CTC CTC GAC GGC CTC GGC GCC CTC 1564 A P DRAR D.G. L 1625 GTC CTG TTC AGC GGC CAG GGT GCC CAA CGT ACG GGC ATG GGC ATG GAG TTG TAC GCC GCC S G Q G A ORT G M G M E 1685 CAC CCC GCC TTC GCG ACG GCC TTC GAC GCC GTC GCC GAA CTG GAC CCC CTC CTC GAC 1744 A A 1745 CGG CCC CTC GCC CAA CTC GTC GCG GCC GGC GAC ACC CTC GAC CGC ACC GTC CAC ACA CAG 1804 LVAAGDTLD R T 1805 CCC GCG CTC TTC GCC GTG GAG GTC GCC CTC CAC CGC CTC GAG TCC TGG GGC GTC ACG 1864 A E. T. H R 1865 CCC GAC CTG CTC GCC GGC CAC TCC GTC GGC GAG ATC AGC GCC GCC CAC GTC GCC GGG GTC 1924 E G A 635 1925 CTG TOG CTG CGC GAC GCC CGC CTC GTC GCG GCG GCC CGC CTC ATG CAG GCG CTC 1984 R D AARLVAA RGRL 1985 CCC GAG GGC GGC GCG ATG GTC GCG GTC GAG GCG AGC GAG GAA GTG CTT CCG CAC CTC 2044 656 P E M V A 7.7 12 A S E E A E 2045 GCG GCA CGC GAG CGG GAG CTC TCC CTC GCG GCC GTG AAC GGC CCC CGC GCG GTC GTC CTC 2104 LAAV N G R A 2105 GCG GCC GAG CGC GCC GTC CTC GAC GTC GCC GAG CTG CTG CGC GAA CAG GGC CGC CGG 2164 A D E A 2165 ACG AAG CGG CTC AGC GTC TCG CAC GCC TTC CAC TCG CCG CTC ATG GAG CCG ATG CTC GAC 2224 2225 GAC TTC CGC CGG GTC GTC GAA GAG CTG GAC TTC CAG GAG CCC CGC GTC GAC GTC GTG TCC 2284 736 D F R v E R E L D F 0 E P R 2285 ACG GTG ACG GGC CTG CCT GTC ACA GCG GGC CAA TGG ACC GAT CCC GAG TAC TGG GTG GAC 2344 A G Q D 2345 CAG GTC CGC AGG CCC GTA CGC TTC CTC GAC GCC GTA CGC ACC CTG GAG GAA TCG GGC GCC 2404 R DAV 2405 GAC ACC TTC CTG GAG CTC GGT CCC GAC GGG GTC TGC TCC GCG ATG GCG GCG GAC TCC GTA 2464 P D G C S A M 215 2465 CGC CAG CAG CAG GCC GCC ACG GCG GTC TCC GCC CTG CGC AAG GGC CGC CCG GAG CCC CAG 2524 A V 2525 TOG CTC CTC GCC GCA CTC ACC ACC GTC TTC GTC CGG GGC CAC GAC GTC GAC TGG ACC GCC 2584 G LAAL Т T R H D 855 2585 GCG CAC GGG AGC ACC GGC ACG GTC AGG GTG CCC CTG CCG ACC TAC GCC TTC CAG CGC GAA 2644 2645 OGC CAG TOG TTC GAC GGC GGC GGG GGG GGG GGC GGC CGG CTC ACG GGG GGC CGA TCG GGC 2704 Т R P L A 2705 ACC GGT GCG GGC ACC GGC CCG GCC GCG GGT GTG ACG TCG GGC GAG GGC GAG GGC GAG GGC G A A G Т S 2765 CAG GGC GCT GCC GCT GGC GGT GAT CGG CCG GCT CGC CAC GAG ACC GAG CGC GTG 2824 G D R н R E 935 2825 CGC GCA CAC GTC GCC GCC GTC CTC GAG TAC GAC GAC CCG ACC CGC GTC GAA CTC GGC CTC 2884 EYDD P т R 955 2885 ACC TTC AAG GAG CTG GGC TTC GAC TCC CTC ATG TCC GTC GAG CTG CGG AAC GCG CTC GTC 2944 D SLMS ELRNA 2945 QAC GAC ACG GGA CTG CGC CTG CCC AGC GGA CTG CTC TTC GAC CAC CCG ACG CCG CGC GCC 3004 R G D н т

FIG. 23B

3005 CTC GCC GCC CAC CTG GGC GAC CTG CTC ACC GGC GGC AGC GGC GAG ACC GGA TCG GCC GAC 3064 A A H L G D L L T G G S G E T G S A 3124 T PAD T A E PI A 1035 3125 GCC TGC CGC TAC CCC GGC GGC GTC ACC TCC CCC GAG GAC CTG TGG CGG CTC GTC GCC GAG 3184 1055 3185 GGG CGC GAC GCC GTC TCG GGG CTG CCC ACC GAC CGC GGC TGG GAC GAG GAC CTC TTC GAC 3244 TDRGWDEDLF SGLP 1075 3245 GCC GAC CCC GAC CGC AGC GGC AAG AGC TCG GTC CGC GAG GGC GGA TTC CTG CAC GAC GCC 3304 G K S S v REGG 1095 3305 GCC CTG TTC GAC GCC GGC TTC TTC GGG ATA TCG CCC CGC GAG GCC CTC GGC ATG GAC CCG R E 3365 CAG CAG CGG CTG CTC CTG GAG ACG GCA TGG GAG GCC GTG GAG CGC GCA GGG CTC GAC CCC 3424 WEA VERA 1135 L E T A G D 3425 GAA GGC CTC AAG GGC AGC CGG ACG GCC GTC TTC GTC GGC GCC ACC GCC CTG GAC TAC GGC 3484 1136 E G L K G S R T A V F V G A T A L D 1155 3485 CCG CGC ATG CAC GAC GGC GCC GAG GGC GTC GAG GGC CAC CTC CTG ACC GGG ACC ACG CCC 3544 1156 P R M H D G A E G V E G H L L T 1175 3545 AGC GTG ATG TCG GGC CGC ATC GCC TAC CAG CTC GGC CTC ACC GGT CCT GCG GTC ACC GTC 3604 IAYQLGL 1195 3605 GAC ACG GCC TGC TCG TCC CTC GTC GCG CTG CAC CTG GCC GTC CGT TCG CTG CGG CAG H A 3665 GGC GAG TCG AGC CTC GCG CTC GCC GGC GGA GCG ACC GTC ATG TCG ACA CCG GGC ATG TTC 3724 v 1216 G E S S L A L A G G A T M S P 1235 3725 GTC GAG TTC TCG CGG CAG CGC GCC CTC GCC GCC GAC GGC CGC TCC AAG GCC TTC TCC GAC 3784 1236 V E F S R Q R G L A A D G R S K A F S 1255 3785 TCC GCC GAC GGC ACC TCC TGG GCC GAG GGC GTC GGC CTC GTC GTC GAG CGG CTC TCG 3844 SWAEGVG L L 1275 3845 GAC GCC GAG CGC AAC GGC CAC CCC GTG CTC GCC GTG ATC CGG GGC AGC GCG GTC AAC CAG 3904 v R 1295 G H P V L A I 3905 GAC GGC GCC TCC AAC GGG CTC ACC GCC CCC AAC GGC CCG TCC CAG CAG CGC GTC ATC CGA 3964 APNG P S 1315 3965 CAG GCC CTG GCC GAC GCC GGG CTC ACC CCG GCC GAC GTC GAC GCC GTC GAG GCG CAC GGT 4024 D A G L P A D 1335 A Т 4025 ACG GGT ACC CGG CTC GGC GAC CCC ATC GAG GCC GAG GCG ATC CTC GGC ACC TAC GGC CGG 4084 1336 T G T R L G D P TEAEAIL 1355 4085 GAC CGG GGC GAG GGC GCT CGG CTC CAG CTC GGC TCG CTG AAG TCG AAC ATC GGC CAC GCG E G A P Q L G S LKS 1375 4145 CAG GCC GCC GCC GCC GTC GCC GCC CTC ATC AAG ATG GTC CTC GCG ATG CGC CAC GGC GTC 4204 V G G L I K M V L A 1395 13760 A A A G 4205 CTG CCC AGG ACG CTC CAC GTG GAC CGG CCC ACC ACC CGC GTC GAC TGG GAG GCC GGC GGC 1396 L P R T L H V D R P T T R V D W E A G G 4264 4324 EEREWPETG 4325 ATC TCC TCC TTC GGC ATC AGC GGC ACC AAC GCC CAC ATC GTG GTC GAA CAG GCC CCG GAA 4384 1455 S G TNAH 4444 TAPEAG 1475 A G Т GEA A 4504 4445 GAC ACC GCC GCC ACC ACG ACG CCG GCC GCG GTC GGC GTC CCC GAA CCC GTA CGC GCC CCC G P E 1495 A A v P 4505 GTC GTG GTC TCC GCG CGG GAC GCC GCC GCC CGC GCC CAG GCC GTT CGG CTG CGG ACC 4564 1515

FIG. 23C

4565 THE CHE GAC GGC CGA COG GAC GHE ACC GHE GGC GAC CHE GGA CGC TOG CHE GCC GCE CGFT 4624 D L G R S 4625 ACC GCC TTC GAG CAC AAG GCC GCC CTC ACC ACC GCC ACC AGG GAC GAG CTG CTC GCC GGG 4684 EHKAAL T T T A D 1555 4685 CTC GAC GCC CTC GGC GGG GAG CAA GCC ACG GGC CTG GTC ACC GGC GAA CCG GCC AGG 4744 1556 L D A L G R G E Q A T G L V 1575 4745 GCC GGA CGC ACG GCC TTC CTG TTC ACC GGC CAG GGA GCG CAG CGC GTC GCC ATG GGC GAG 4804 TGQGAQR 1595 4805 GAA CTG CGC GCC GCG CAC CCC GTG TTC GCC GCC GCC CTC GAC ACC GTG TAC GCG GCC CTC 4864 A A L 4865 CAC COT CAC CTC GAC COG CCG CTG COG GAG ATC GTC GCC GCC GGG GAG GAG CTG GAC CTC 4924 D R R E I V E E L 4925 ACC GCG TAC ACC CAG CCC GCC CTC TTC GCC TTC GAG GTG GCG CTG TTC CGC CTC GAA 4984 LFAFE v A 4985 CAC CAC GGC CTC GTC CCC GAC CTG CTC ACC GGC CAC TCC GTC GGC GAG ATC GCC GCC GCG 5044 1656 H H G L V P D LLTGH 1675 5045 CAC GTC GCC GGT GTC CTC TCC CTC GAC GAC GCC GCA CGT CTC GTC ACC GCC CGC GGC CGG 5104 LDDAARL 1695 5105 CTC ATG CAG TCG GCC CGC GAG GGC GGC GCG ATG ATC GCC GTG CAG GCG GGC GAG GCC GAG 5164 EGGAMI 5165 GTC GTC GAG TCC CTG AAG GGC TAC GAG GGC AGG GTC GCC GTC GCC GTC AAC GGA CCC 5224 G R V L.KG E 5225 ACC GCC GTG GTC GTC TCC GGC GAC GCG GAC GCC GCC GAG GAG ATC CGC GCC GTA TGG GCG 5284 GDADAA EEIR 5285 GGA CGC GGC CGG CGC ACC CGC AGG CTG CGC GTC AGC CAC GCC TTC CAC TCC CCG CAC ATG 5344 RRLRVS H A 1775 5345 GAC GAC GTC CTC GAC GAG TTC CTC CGG GTC GCC GAG GGC CTG ACC TTC GAG GAG CCG CGG 5404 1795 5405 ATC CCC GTC GTC TCC ACG GTC ACC GGC GCC GTC GTC ACG TCC GGC GAG CTC ACC TCG CCC 5464 A L 5465 GCG TAC TGG GTC GAC CAG ATC CGG CGG CCC GTG CGC TTC CTG GAC GCC GTC CGC ACC CTG 5524 V 5525 GCC GCC CAG GAC GCG ACC GTC CTC GTC GAG ATC GGC CCC GAC GCC GTC CTC ACG GCA CTC 5584 1836 A A Q D A T I. E I G PDAV 1855 5585 GCC GAG GAG GCT CTC GCG CCC GGC ACG GAC GCC CCG GAC GCC CGG GAC GTC ACG GTC GTC 5644 1875 5645 CCG CTG CTG CGC GCG GGG CGC CCC GAG CCC GAG ACC CTC GCC GCC GCT CTC GCG ACC GCC 5704 RAGR E T 1895 5705 CAT GTC CAC GGC GCA CCC TTG GAC CGG GCG TCG TTC TTC CCG GAC GGG CGC CGC ACG GAC 5764 APLDRA 5765 CTG CCC ACG TAC GCC TTC CGG CGC GAG CAC TAC TGG CTG ACG CCC GAG GCC CGT ACG GAC 5824 R R E H 1935 5825 OCC CGC GCA CTC GGC TTC GAC CCG GCG CGG CÄC CCG CTG CTG ACG ACC ACG GTC GAG GTC 5884 ARHPLL 1955 5885 GCC GGC GGC GAC GGC GTC CTG CTG ACC GGC CGT CTC TCC CTG ACC GAC CAG CCC TGG CTG 5944 Т GRL 1975 5945 GCC GAC CAC ATG GTC AAC GGC GCC GTC CTG TTG CCG GCC ACC GCC TTC CTG GAG CTC GCC D H M V N G A LL 1995 6005 CTC GCG GCG GGC GAC CAC GTC GGG GCG GTC CGG GTG GAG GAA CTC ACC CTC GAA GCG CCG 6064 1996 L A A G D v A V v Н G R E т 6065 CTC GTC CTG CCC GAG CGG GGC GCC GTC CGC ATC CAG GTC GGC GTG AGC GGC GAC GGC GAG 2016 L V L P E R G A V R I O V G V S G D G E 6124 2035

FIG. 23D

6184 2055 6185 GAC GCG CCC CGG GAG TGG ACC CGC CAT GTC TCC GGC GTA CTC GGC GAA GGG GAC CCG GCC 6244 R H 2075 6245 ACG GAG TOG GAC CAC CCC GGC ACC GAC GGG GAC GGT TCA GCG GCC TGG CCG CCT GCG GCC 6304 D G D G G A 2095 6305 GCG ACC GCC ACA CCC CTC GAC GGC GTC TAC GAC CGG CTC GCG GAG CTC GGC TAC GGA TAC 6364 D G V R LAEL 2115 6365 GGT CCG GCC TTC CAG GGC CTG ACG GGG CTG TGG CGC GAC GGC GCC GAC ACG CTC GCC GAG 6424 Q G L T G L W R D G A D 6425 ATC CGG CTG CCC GCG GCG CAG CAC GAG AGC GCG GGG CTC TTC GGC GTA CAC CCG GCG CTG 6484 о н A G G 2155 6485 CTC GAC GCG GCG CTC CAC CCG ATC GTC CTG GAG GGC AAC TCA GCT GCC GGT GCC TGT GAC 6544 v L H LEG 6545 GCC CAT ACC CAC GCG ACC CAC CGG ATC CGG CTG CCG TTC GCG TGG GCG GCG GTG ACC CTC 6604 TDRI R L P 2195 6605 CAC GCC GAA GGG GCC ACC GGG CTC CGC GTA CGG ATC ACA CCC ACC GGC CCG GAC ACG GTC 6664 2196 H A E G A L R R 2215 6665 ACC CTC CGC CTC ACC GAC ACC ACC GGT GCG CCC GTG GCC ACC GTG GAG TCC CTG ACC CTG 6724 T T T G D A P V A T v E 6725 CGC GCG GTG GCG AAG GAC CGG CTG GGC ACC ACC GCC GGG CGC GTC GAC GCC CTG TTC 2236 R A V A K D R L G T T A G R V D D A L F 6784 2255 6785 ACG GTC GTG TGG ACG GAG ACC GGC ACA CCC GAA CCC GCA GGG CGC GGA GCC GTG GAG GTC 6844 2275 6845 GAG GAA CTC GTC GAC CTC GGC GGC CTC GGC GAC CTC GTG GAG CTC GGC GCC GCG GAC GTC 6904 D A D E L 2295 6905 GTC CTC CGG GCC GAC CGC TGG ACG CTC GAC GGG GAC CCG TCC GCC GCC GCC ACA GCC 6964 DRWT D P S A 2315 6965 GTC CGG CGC ACC CTC GCC ATC GTC CAG GAG TTC CTG TCC GAG CCG CGC TTC GAC GGC TCG 7024 2316 V R R T L A Т Q E F L S E P R F D 2335 G 7025 CHA CTG GTG TGC GTC ACC AGG GGC GCG GTC GCC GCA CTC CCC GGC GAG GAC GTC ACC TCC 7084 2355 7085 CTC GCC ACC GGC CCC CTC TGG GGC CTC GTC CGC TCC GCC CAG TCC GAG AAC CCG GGA CGC 7144 R 2375 7145 CTG TTC CTC CTG GAC CTG GGT GAA GGC GAA GGC GAG CGC GAC GGA GCC GAG CTG ATC 7204 D L G E G EGERD 2395 7205 CGC GCG GCC ACG GCC GGG GAC GAG CCG CAG CTC GCG GCA CGG GCA CGG CCA CTC GCG 7264 Q L A A R 2396 R A A GDE A T D G R L 2415 7265 CCG AGG CTG GCC CGT ACC GCC GCC CTT TCG AGT GAG GAC ACC GCC GGC GGC GCC GAC CGT 7324 2416 P R L A R T A A L SSEDT A G G 2435 7325 TTC GGC CCC GAC GGC ACC GTC CTC GTC, ACC GGG GGC ACC GGA GGC CTC GGA GCG CTC CTC 7384 D T T G G G G G. L 2455 7385 GCC CGC CAC CTC GTG GAG CGT CAC GGG GTG CGC CGG CTG CTG CTG GTG AGC CGC CGC GGG 7444 H R 2475 7445 GCC GAC GCC CCC GGC GCC GAC CTG GGC GAG GAC CTC GCG GGC CTC GGC GCG GAG GTG 7504 D G E D A G A A A G G 2495 7505 GCG TTC GCC GCC GCC GAC GCC GAC CGC GAG AGC CTG GCG CGG GCG ATC GCC ACC GTG 7564 DAADRES 2496 A F A A LARA 2515 7565 CCC GCC GAG CAT CCG CTG ACG GCC GTC GTG CAC ACG GCG GGA GTC GTC GAC GAC GCG ACG 7624 v 2535 7625 GTG GAG GCG CTC ACA CCG GAA CGG CTG GAC GCG GTA CTG CGC CCG AAG GTC GAC GCC GCG 7684 ERLDAV R P

FIG. 23E

7685 TOG AAC CTG CAC GAG CTC ACC AAG GAC CTG COG CTC GAC GCC TTC GTC CTC TCC TCC 7744 2556 W N L H E L T K D L R L D A F v 2575 7745 GTC TCC GGC ATC GTC GGC ACC GCC GGC CAG GCC AAC TAC GCG GCG GCC AAC ACG GGC CTC 7804 V G T A G Q A N Y A A A N 2595 7805 GAC GCC CTC GCC GCC CAC CGC GCC GCC ACG GGC CTG GCC GCC ACG TCG CTC GCC TGG GGC 7864 2596 D A L A A H R A A TGLAATSLAW 2615 7865 CTC TGG GAC GGC ACG CAC GGC ATG GGC GGC ACG CTC GGC GCC GCC GAC CTC GCC CGC TGG 7924 2616 L W D G T H G M G G T L G A D A 2635 7925 AGC CGG GCC GGA ATC ACC CCG CTC ACC CCG CTG CAG GGC CTC GCG CTC TTC GAC GCC GCG 2655 7985 GTC GCC AGG GAC GAC GCC CTC CTC GTA CCC GGC GGG CTC CGT CCC ACC GCC CAC CGG GGC 8044 R D V G DALL P A L R P T 8045 ACG GAC GGA CAG CCT CCT GCG CTG TGG CGC GGC CTC GTC CGG GCG CGC CCC CGT GCC 8104 PALWRG R A R P 2695 8164 A E A A D T T G G W L S G L 2715 8165 CAG TCC CCC GAG GAG CGC CGC AGC ACA GCC GTC ACG CTC GTG ACG GGT GTC GTC GCG GAC 8224 E RRS T. A т L v 2735 8225 GTC CTC GGG CAC GCC GAC TCC GCC GCG GTC GGG GCG GAG CGG TCC TTC AAG GAC CTC GGC A D н A G AERSF A 8285 TTC GAC TCC CTG GCC GGG GTG GAG CTC CGC AAC CGG CTG AAC GCC GCC ACC GGC CTG CGG 8344 A G v ELRNRLNAA 8345 CTC CCC GCG ACC ACG GTC TTC GAC CAT CCC TCG CCG GCC GCG CTC GCG TCC CAT CTC CTC 8404 FDHP S S 2795 8405 GCC CAG GTG CCC GGG TTG AAG GAG GGG ACG GCG ACC GCG ACC GTC GTG GCC GAG CGG 8464 GLKEG TAATAT 2815 8465 GGC GCT TCC TTC GGT GAC CGT GCG ACC GAC GAC GAT CCG ATC GCG ATC GTG GGC ATG GCA 8524 I GDRATDDD P 2835 8525 TGC CGC TAT CCG GGT GGT GTG TCG TCG CCG GAG GAC CTG TGG CGG CTG GTG GCC GAG GGG 8584 G v 8585 ACG GAC GCG ATC AGC GAG TTC CCC GTC AAC CGC GGC TGG GAC CTG GAG AGC CTC TAC GAC 8644 2856 T D A I S E F V P NRGWD E S 2875 8645 CCG GAT CCC GAG TCG AAG GGC ACC ACG TAC TGC CGG GAG GGC GGG TTC CTG GAA GGC GCC 8704 2895 8705 GGT GAC TTC GAC GCC GCC TTC TTC GGC ATC TCG CCG CGC GAG GCC CTG GTG ATG GAC CCG 8764 D A A F GISPRE 2915 A L 8765 CAG CAG CGC CTG CTG CAG GTG TCC TGG GAG GCG CTG GAA CGC GCG GGC ATC GAC CCG W E A L R 2935 A 8825 TCC TCG CTG CGC GGC AGC CGC GGT GGT GTC TAC GTG GGC GCC GCG CAC GGC TCG TAC GCC 8884 G S G S L R G v Y W G 2955 R н 8885 TCC GAT CCC CGG CTG GTG CCC GAG GGC TCG GAG GGC TAT CTG CTG ACC GGC AGC GCC GAC 8944 2975 8945 GCG GTG ATG TCC GGC CGC ATC TCC TAC GCG CTC GGT CTC GAA GGA CCG TCC ATG ACG GTG 9004 S Y A L G L E G P 2995 9005 GAG ACG GCC TGC TCC TCC CTG GTG GCG CTG CAT CTG GCG GTA CGG GCG CTG CGG CAC 9064 v L A S S S L ALH Α VRA 3015 9065 GCC GAG TGC GCG CTC GCG CTG GCC GCC GCG GTG GCC GTG ATG GCC GAT CCC GCG GCG TTC 9124 G A Α G G V A V A D 3035 9125 GTG GAG TTC TCC CGG CAG AAG GGG CTG GCC GCC GAC GGC CGC TGC AAG GCG TTC TCG GCC 9184 KGLA R O A D G R 3055 9185 GCC GCC GAC GGC ACC GGC TGG GCC GAG GGC GTC GGC GTC GTC CTG GAG CGG CTG TCG 9244 3075

FIG. 23F

9245 GAC GCG CGC CGC GCG CAC ACG GTC CTC GGC CTC GTC ACC GGC ACC GCG GTC AAC CAG 9304 3076 D A R R A G H T V L G L V T G T A V N Q 3095 9305 GAC GGT GCC TCC AAC GGG CTG ACC GCG CCC AAC GGC CCA GCC CAG CAA CGC GTC ATC GCC 9364 3096 D G A S N G L T A P N G P A Q Q 3115 9365 CAG GCG CTC GCC GAC GCC GGG CTG TCC CCG GAG GAC GTG GAC GCG GTC GAG GCG CAC GGC 9424 מ 3116 E A L A D A G L S P E D Α 3135 9425 ACC GGC ACC CGG CTC GGC GAC CCC ATC GAG GCC GGG GCG CTC GCC GCC TCC GGA CGG 9484 3155 RLGDP IEA G A L 9485 AAC COT TOO GGC CAC CAC CCG CTG TGG CTC GGC TCG CTG AAG TCC AAC ATC GGG CAT GCC 9544 W 3175 D H P 9604 9545 CAG GCC GCC GCT GTC GGC GGC GTC ATC AAG ATG CTC CAG GCG CTG CGG CAC GGC TTG v KMLQALR 3195 31760 A A A G V G G I 9605 CTG CCC CGC ACC CTC CAC GCC GAC GAG CCG ACC CCG CAT GCC GAC TGG AGC TCC GGC CGG 9664 T PHADWS 3215 TLHADEP 3196 L R 9665 GTA COG CTG CTC ACC TCC GAG GTG CCG TGG CAG CGG ACC GGC CCG CCG CCG ACC GGG 9724 RLLTSEV PWQRTGRP 3235 9725 GTG TCC GCC TTC GGC GTC GGC ACC AAT GCC CAT GTC GTC CTC GAA GAG GCA CCC GCC G G G Т N A H v L E А 9785 CCG CCC GCG CCG GAA CCG GCC GGG GAG GCC CCC GGC GCC TCC CGC GCC GCA GAA GGG GCG 9844 P G R A E P G E A 9845 GAA GGG CCC CTG GCC TGG GTG GTC TCC GGA CGC GAC GAG CCG GCC CTG CGG TCC CAG GCC 9904 SGRDEP A R S 3295 3276 E G P L A 9964 9905 CGC CGC CTC CGC CAC CTC TCC CGC ACC CCC GGG GCC CGC CGT CAC ATC GCC TTC 3315 R D H L S RTPGARPRD 3296 R 9965 TCC CTC GCC GCC ACG CGC GCA GCC TTT GAC CAC CGC GCC GTG CTG ATC GGC TCG GAC GGG 10024 3316 S L A A T R A A F D H R A V L G 3335 10025 GCC GAA CTC GCC GCC GCC CTG GAC GCG TTG GCC GAA GGA CGC GAC GGT CCG GCG GTG GTG 10084 R D G A LDALAEG 10085 CGC GGA GTC CGC GAC CGG GAC GGC AGG ATG GCC TTC CTC TTC ACC GGG CAG GGC AGC CAG 10144 R D G R M A G 3375 10145 CGC GCC GGG ATG GCC CAC GAC CTG CAT GCC GCC CAT ACC TTC TTC GCG TCC GCC CTC GAC 10204 3395 3376 R A G M A H D L H A A H т F A 10264 TDRLDPLLGRPLGAL 3415 3396 E V 10265 CCA CCC GCC TCG CCC GAA GCG GCA CTC CTG GAC CCG ACC GAG TAC ACC CAG CCG GCG CTC 10324 LDRTEY 3435 3416 R P G S P E A A L Т 0 10325 TTC GCC GTC GAG GTG GCG CTC CAC CGG CTG CTG GAG CAC TGG GGG ATG CGC CCC GAC CTG 10384 3455 LHRLL E H W G M R D E A 10385 CTG CTG GGG CAC TCG GTG GGC GAA CTG GCG GCC GCC CAC GTC GCG GGT GTG CTC GAT CTC 10444 3475 A A H A G L D v G E L A 10445 CAC GAC GCC TGC GCC CTG GTG GCC GCC GGC AGG CTG ATG CAG GGC CTG CCG GCC GGC 10504 DACALVAARGRLMQR 3495 3476 D 10505 GGC GCG ATG GTC TCC GTG CGG GCC GGC GAG GAC GAG GTC CGC GCA CTG CTG GCC GGC CGC 10564 3515 E D E R A V R A G 10565 GAG GAC GCC GTC TGC GTC GCC GCG GTG AAC GGC CCC CGG TCG GTG ATC TCC GGC GCG 10624 P R S 3535 A A v N G 10625 GAG GAA GCG GTG GCC GAG GCG GCG GCG CAG CTC GCC GGA CGA GCC CGC CGC ACC ACG CGG 10684 R 3555 G R GRR EAAAQLA Α 10685 CTC CGC GTC GCG CAC GCC TTC CAC TCA CCC CTG ATG GAC GGC ATG CTC GCC GGA TTC CGG 10744 T. M D G S Н A н 10745 GAG GTC GCC GCC GCC CTG CGC TAC CGG GAA CCG GAG CTG ACG GTC GTC TCC ACG GTC ACG 10804 3595 Y 3576 E V A AGL R

FIG. 23G

10805 GGG CGG CCC GCC CGC GGT GAA CTC ACC GGC CCC GAC TAC TGG GTG GCC CAG GTC CGT 10864 3596 G R GELTG D Y PA R ₽ P W V. A 3615 10865 CAG CCC GTG CGC TTC GCG GAC GCG GTC CGC ACG GCA CAC CGC CTC GGA GCC CGC ACC TTC 10924 3616 E P V R T A A D A V R н R L G $\mathbf{R} \cdot \mathbf{T}$ 3635 10925 CTG GAG ACC GGC CCG GAC GGC GTG CTG TGC GGC ATG GCA GAG GAG TGC CTG GAG GAC GAC 10984 GVLCGM AEE 3655 10985 ACC GTG GCC CTG CTG CCG GCG ATC CAC AAG CCC GGC ACC GCG CCG CAC GGT CCG GCC GCT 11044 T A P AIHKPG P 3675 11045 CCC GGC GCG CTG CGG GCC GCC GCC GCC GCG TAC GGC CGG GGC GCC CGG GTG GAC TGG GCC 11104 GAL A R A A A A Y G RGA 11164 A R R 3715 11165 TTC CGG CAC CGC CGC TAC TGG CTC GCC CCG GGC CGC GCG GAC ACC GAC GAC TGG ATG 11224 R'H R Y W L P G R R A A A D D 3735 11225 TAC CGG ATC GGC TGG GAC CGG CTG CCG GCT GTG ACC GGC GGG GCC CGG ACC GCC CGC 11284 3736 Y R I G WDRLPAV Т GGA R 3755 A 11285 TGG CTG GTG ATC CAC CCC GAC AGC CGG CGC TGC CGG GAG CTG TCC GGC CAC GCC GAA CGC 11344 L·V I H DSPRC RELS G H A 3775 11345 GCG CTG CGC GCC GCG GCC GCG AGC CCC GTA CCG CTG CCC GTG GAC GCT CCG GCC GCC GAC 11404 3776 A L R A A G A S P v P P D 11405 CGG GCG TCC TTC GCG GCA CTG CTG CGC TCC GCC ACC GGA CCT GAC ACA CGA GGT GAC ACA 11464 R A 3815 11465 GCC GCG CCC GTG GCC GGT GTG CTG TCG CTG TCC GAG GAG GAT CGG CCC CAT CGC CAG 11524 S 3816 A A P A G v L S L L E E D R P 3835 11525 CAC GCC CCG GTA CCC GCC GGG GTC CTG GCG ACG CTG TCC CTG ATG CAG GCT ATG GAG GAG 11584 L A T SLMOAM 3855 11644 v E A R v R 3875 W S A A v 11704 A G A A L W G L G R Α 3895 11705 CGC CCC ACC CGG TGG GGC GGT CTC GTG GAC CTG CCC GCC TCG CCC GGT GCG GCG CAC TGG 11764 G G v D P Α 3915 11765 GCG GCC GCC GTG GAA CGG CTC GCC GGT CCC GAG GAC CAG ATC GCC GTG GGC GCG TCC GGC 11824 VERLA E D QIA G R A 3935 11825 AGT TOG OGC COG CGC CTC ACC AGG CTG CCG CGC GAC GGC GGC CGG ACG GCC GCA CCC 11884 GRR L T R P R D G G G 3955 11985 GCG TAC CGG CCG CGC GGC ACG GTG CTC GTC ACC GGT GGC ACC GGC GCG CTC GGC GGG CAT 11944 3956 A T v V R R G L T G G т G A G 3975 11945 CTC GCC CGC TGG CTC GCC GCG GCG GCC GAA CAC CTG GCG CTC ACC AGC CGC CGG GGC 12004 WLAAA GAEH L A L 3995 12005 CCG GAC GCG CCC GCC GCC GCA CTC GAG GCC GAA CTC CTC CTC CTC GCC GCC AAG GTG 12064 A E 3996 P D A Α A G L E 4015 12065 ACG TTC GCC GCC TGC GAC ACC GCC GAC CGC GAC GGC CTC GCC CGG GTC CTG CGG GCG ATA 12124 C D DRDGLA 4035 12125 CCC GAG GAC ACC CCC CTC ACC GCC GTG TTC CAC GCC GCC GCC GTA CCC CAG GTC ACC CCC 12184 V F Т A HAA G P 0 4055 12185 CTG TCC CGT ACC TCG CCC QAG CAC TTC GCC QAC GTG TAC GCG GGC AAG GCG GCG GGC GCC 12244 E H F A D v A K A 4075 12245 GCG CAC CTG GAC GAA CTG ACC CGC GAA CTC GGC GCC GGA CTC GAC GCG TTC GTC CTC TAC 12304 4076 A H I. D E L Т R E G A G L D A 4095 L 12305 TCC TCC GGC GCC GGC GTC TGG GGC AGC GCC GGC CAG GGT GCC TAC GCC GCC GCC AAC GCC 12364 4096 S S SAGQGAY 4115

FIG. 23H

12365 GCC CTG GAC GCG CTC GCC CGG CGC CGT GCG GCG GAC GGA CTC CCC GCC ACC TCC ATC GCC 12424 4116 A L D A L A R R R A A D G L P A T 4135 12425 TOG GGC GTG TOG GGC GGC GGC GGT ATG GGG GCC GAC GAG GCG GGC GCG GAG TAT CTG GGC 12484 G G G G G M G A D E Α 4155 12485 CGC CGC GGT ATG CGC CCC ATG GCA CCG GTC TCC GCG CTC CGG GCG ATG GCC ACC GCC ATC 12544 Α M R P M A P R R G Α M Α 4175 12545 GCC TCC GGG GAA CCC TGC CCC ACC GTC ACC CAC ACC GAC TGG GAG CGC TTC GGC GAG GGC 12604 T D 4176 A S G E P C P T V T Н W E R 4195 12605 TTC ACC GCC TTC CGG CCC AGC .CCT CTG ATC GCG GGG CTC GGC ACG CCG GGC GGC GGC CGG 12664 4196 F т R S I Α G L G Α 4215 12665 GCG GCG GAG ACC CCC GAG GAG GGG AAC GCC ACC GCT GCG GCG GAC CTC ACC GCC CTG CCG 12724 PEEGN ATAAADL Е Т T 4235 A 12725 CCC GCC GAA CTC CGC ACC GCG CTG CGC GAG CTG GTG CGA GCC CGG ACC GCG GCG CTC 12784 R A R 4236 P A E L R T A L R E L V 4255 12785 GGC CTC GAC GAC CCG GCC GAG GTC GCC GAG GGC GAA CCG TTC CCC GCC ATG GGC TTC GAC 12844 D D P A E v A EGERF P Α M G 4275 12845 TCC CTG GCC ACC GTA CGG CTG CGC CGC GGA CTC GCC TCG GCC ACG GGC CTC GAC CTG CCC 12904 R L R R G L A , S A T а т 4295 12905 CCC GAT CTG CTC TTC GAC CGG GAC ACC CCG GCC GCC GCC GCC CAC CTG GCC GAA CTG 12964 Т P A L D R D A A. A Н 4315 13024 TARDHG G G P G A 4335 4316 L A P A P 13025 GCC GGA AGC GGC CTG CCG GCC CTC TAC CGG GAG GCC GTC CGC ACC GGC CGG GCC GCG GAA 13084 4336 A G S G L P A L Y R E A v R 13085 ATG GCC GAA CTG CTC GCC GCC GCT TCC CGG TTC CGC CCC GCC TTC GGG ACG GCG GAC CGG 13144 R F 4356 M A E L A A Α S R P Α G 4375 13145 CAG CCG GTG GCC CTC GTG CCG CTG GCC GAC GGC GAG GAC ACC GGG CTC CCG CTG CTC 13204 4376 O P V A L V PLADGA E D T G 4395 13205 GTG GGC TGC GCC GGG ACG GCG GTG GCC TCC GGC CCG GTG GAG TTC ACC GCC TTC GCC GGA 13264 Α 4396 V G A 4415 13265 GCG CTG GCG GAC CTC CCG GCG GCC GCC CCG ATG GCC GCG CTG CCG CAG CCC GGC TTT CTG 13324 AAAPMA 13325 CCG GGA GAA CGA GTC CCG GCC ACC CCG GAG GCA TTG TTC GAG GCC CAG GCG GAA GCG CTG 13384 F E 4455 4436 P G E R PATPEA T. A O A F. 13385 CTG CGC TAC GCG GCC GGC CGG CCC TTC GTG CTG CTG GGG CAC TCC GCC GGC GCC AAC ATG 13444 R A A G R v G H A G 13445 GCC CAC GCC CTG ACC CGT CAT CTG GAG GCG AAC GGT GGC GGC CCC GCA GGG CTG GTG CTC 13504 T R H E A NGGGP A G 4495 H A L L 13505 ATG GAC ATC TAC ACC CCC GCC GAC CCC GGC GCG ATG GGC GTC TGG CGG AAC GAC ATG TTC 13564 P G A M G V W R N 13565 CAG TOP GTC TOG COG CGC TCG GAC ATC CCC CCG GAC GAC CAC CGC CTC ACG GCC ATG GGC 13624 R R D D D H R L 4535 S 13625 GCC TAC CAC CGG CTG CTT CTC GAC TGG TCG CCC ACC CCC GTC CGC GCC CCC GTA CTG CAT 13684 L D W S P Т P v A 4555 HRLL 13685 CTG CGC GCC GCG GAA CCC ATG GGC GAC TOG CCA CCC GGG GAC ACC GGC TGG CAG TCC CAC 13744 G D M G D W P P G W 4575 4556 L R A A E P 13745 TOG GAC GGC GCG CAC ACC ACC GCC GGC ATC CCC GGA AAC CAC TTC ACG ATG ATG ACC GAA 13804 GIP G N н 4595 Т T A 4576 W D G A н 13805 CAC GCC TCC GCC GCC GCC CGG CTC GTG CAC GGC TGG CTC GCG GAA CGG ACC CCG TCC GGG 13864 V R 4615 4596 H A S A A A R L H G W L A E т 13865 CAG GGC GGG TCA CCG TCC CGC GCG GGG AGA GAG GAG AGG CCG TGA ACACGGCAGCCGGCCC 13928 S R A

FIG. 231

14009 CCCGGGCCGCACAGTGGTTCGCCGGCAACCAGGGAACCCCTACGGG ATG ATC CTG CGC GCC GCC ACC GCC 14079 MILRAG 14080 GAC CCG GCA CCG TAC GAG GAA GAG ATC CCC GGG TAC CGA GCT CGA ATT CIT AAT TAA GGAG 14140 E E E I P G Y 14141 GTCGTAG ATG AGT AAC AAG AAC AAC GAT GAG CTG CAG CGG CAG GCC TCG GAA AAC ACC CTG 14201 18 NKNND ELQ RQASEN 14202 GGG CTG AAC CCG GTC ATC GGT ATC CGC CGC AAA GAC CTG TTG AGC TCG GCA CGC ACC GTG 14261 IGIRRKD LL A R 38 14262 CTG CGC CAG GCC GTG CGC CAA CCG CTG CAC AGC GCC AAG CAT GTG GCC CAC TTT GGC CTG 14321 S v A R 0 P L H A K H A Н 58 14322 GAG CTG AAG AAC GTG CTG GGC AAG TCC AGC CTT GCC CCG GAA AGC GAC GAC CGT CGC 14381 K N K S S APE VLLG 14382 TTC AAT GAC CCG GCA TGG AGC AAC AAC CCA CTT TAC CGC CGC TAC CTG CAA ACC TAT CTG 14441 D P A W S N N P L Y RRYLQ 14442 GCC TGG CGC AAG GAG CTG CAG GAC TGG ATC GGC AAC AGC GAC CTG TCG CCC CAG GAC ATC 14501 WIGN S D L RKELOD 14502 AGC CGC CGC CAG TTC GTC ATC AAC CTG ATG ACC GAA GCC ATG GCT CCG ACC AAC ACC CTG 14561 I N T E A M A G O F L M P 138 14562 TCC AAC CCG GCA GCA GTC AAA CGC TTC TTC GAA ACC GGC GGC AAG AGC CTG CTC GAT GGC 14621 139 S N P A A VKRFF E T GGK 14681 14682 GAC GCC TTC GAG GTG GGC AAG AAC CTG GGC ACC AGT GAA GGC GCC GTG GTG TAC CGC AAC 14741 179 D A F V G KNLGTSEGA 198 14742 GAT GTG CTG GAG CTG ATC CAG TAC AAG CCC ATC ACC GAG CAG GTG CAT GCC CGC CCG CTG 14801 Q YKP I T E O н R 218 14802 CTG GTG GTG CCG CCG CAG ATC AAC AAG TTC TAC GTA TTC GAC CTG AGC CCG GAA AAG AGC 14861 I N K F 219 L V 0 14862 CTG GCA CGC TAC TGC CTG CGC TCG CAG CAG ACC TTC ATC ATC AGC TGG CGC AAC CCG 14921 A R C R S T F Ι I S R 258 L Q Q Q14922 ACC AAA GCC CAG CGC GAA TGG GGC CTG TCC ACC TAC ATC GAC GCG CTC AAG GAG GCG GTC 14981 259 T K A Q R E W G L S 14982 GAC GCG GTG CTG GCG ATT ACC GGC AGC AAG GAC CTG AAC ATG CTC GGT GCC TGC TCC GGC 15041 G G S K D L N M L 298 Т 15042 GGC ATC ACC TGC ACG GCA TTG GTC GGC CAC TAT GCC GCC CTC GGC GAA AAC AAG GTC AAT 15101 A V G H Y A L G 318 T A 15102 GCC CTG ACC CTG CTG GTC AGC GTG CTG GAC ACC ACC ATG GAC AAC CAG GTC GCC CTG TTC 15161 V v D Т Т M D N Q 338 15162 GTC GAC GAG CAG ACT TTG GAG GCC GCC AAG CGC CAC TCC TAC CAG GCC GGT GTG CTC GAA 15221 Y 358 Q T L E A A KRHS E Q A G 15222 GGC AGC GAG ATG GCC AAG GTG TTC GCC TGG ATG CGC CCC AAC GAC CTG ATC TGG AAC TAC 15281 D AWMRPN 15282 TGG GTC AAC AAC TAC CTG CTC GGC AAC GAG CCG CCG GTG TTC GAC ATC CTG TTC TGG AAC 15341 G N E P 15342 AAC GAC ACG ACG CCG CCG GCC GCC TTC CAC GGC GAC CTG ATC GAA ATG TTC AAG AGC 15401 . R L PAAFHGDL I E 15402 AAC CCG CTG ACC CCG GAC GCC CTG GAG GTT TGC GGC ACT CCG ATC GAC CTG AAA CAG 15461 E V C G 419 N T P 438 D Α 15462 GTC AAA TGC GAC ATC TAC AGC CTT GCC GGC ACC AAC GAC CAC ATC ACC CCG TGG CAG TCA 15521 S

FIG. 23J

15522 459										GGC G						CTG L		AAC N .		GGC G	15581 478
15582 479										GGC G								ACC T		GCC A	15641 498
15642 499	GAT D	CGC R	CCG P	GGT G	D Gýc	CCG P	GTG V	GCC A	TGG W	CAG Q	GAA E	AAC N	GCC A	ACC T	AAG K	CAT H	GCC A	GAC D	TCC S	TGG W	15701 518
										GAG E											15761 538
15762 539	OGC R	CTG L	G G G	aac N	CGT R	GCC A	TAT Y	GCC A	GCT A	GGC G	gag E	GCA A	TCC S	CCG P	GGC G	ACC T	TAC Y	V GTT	CÀC H	GAG. E	15821 558
15822 559		TGA *	GCT	GCAG	ccc	GTGG	CCAC	CTGC	GGGA	cccc	ACGG	IGTI	GAAT	TC						•	15872 560

TITLE: PROMOTER FOR METHYMYCIN AND PIKROMYCIN INVENTORS NAME: David H. Sherman et al.

SERIAL NO.: 09/988,384

REPLACEMENT SHEET

CATENTA	SP 2 4 2004 55
	RACEMARK DEED

Φ
ပ
C
9
2
ō
æ
0)
ഗ
Ÿ
ᅐ
室
☴

1234567890 1234567890 1234567890 12345678	234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
Bamki V GGAICCGACC GIGGGIGIGA AICICCGGGI GCICGCCICG ICCIGCCCGG TIACCIGICC GCCICCCGCI CCAGACCAGC GGGAGGCGGA CAGGGGCAIG	IGGGTGTGA	ATCTCCGGGT	GCTCGCCTCG	TCCTGCCCCG	TPACCIGICC	GCCTCCCGCT	CCAGACCAGC	GGGAGGCGGA	CAGGGGCATG	100
Sphi ▼ cccccccccc cctaacsc ccateccsc iccatacsac sagecticscs cscentsscs secentssic iscosses siscesses iscassas	3CTAACGGC_	CCGTGCGGCG	TCCGTACGAC	GAGCCTCGCG	CGCCCTGGCG	GCCCTTGGTC	TGCCGGACCT	<u>GTGCGCGGGG.</u>	TGCGCAGGGT	200
BBLXI TRECRECCEC GEGIGGGGC GIAICIGCGG CICCCGGGCA CGGCGGCCT GCTCGICICC GAGICATAGI CCCTGCCGCC GGCGCCACCG CCCTGGCCCG	CGTGGGGCC	GTATCTGCGG	CTCCCGGGGCA	CGGCGGCCCT	GCTCGTCTCC	GAGTCATAGT	CCCTGCCGCC	GGCGCCACCG	BatxI CCCTGGCCCG	41/164 e
Sphi ▼ GCANGGGGGT GCGGGGGCC CCCGGCGCGT AACTCGGCTG GGAGGCCTGG AAAAGGGCGA TCCATTGGGT GAGCGTGAGG TCCTTCGGCA GTCCGCCGTC	້ວວອວອອອວວ	CCCGGCGCGT	AACTCGGCTG	GGAGGCCTGG	AAAAGGGCGA	TCCATTGGGT	GAGCGTGAGG	rccrrcssca	GTCCGCCGTC	400
ECORI Apol ▼ CHARATTCCG TGGCGGTCGG CGAGGGAACG GTAGGTCCGC TTGGGGATGT GGCGCCGGAG GATCTCCGCG AGGCCCCGTC CGGGGCCGGT GAAGACGGCT	eecearcea.	CGAGGGBACG	GTAGGTCCGC	TTGGGGATGT	GGCGCCGGAG	GATCTCCGCG	AGGCCCCGTC	CGGGGCCGGT	GAAGACGGCT	. 000

FIG. 31A

REPLACEMENT SHEET

pikPKS Sequence

FIG. 31B

AGTICIGAGIC CATCICCACG ACCGICACGG IGICGAAGGA GCGCACCAAC ICCICGGITA ICGCGCCCIT ICCGGGGCCG AITICGAGAA CGIICCIACC

RFPI	ACF	AFNT	SHEET	
REPL	AUEN	ACIN I	SULLI	

	1100	1200	1300	1400	1500
1234567890	נפכבכבב	acgaccggag	gtccggagga	gctccaggag	atcggctgtt
	GICCCCCICG ACATGCGIGA CGAGAITIGGG CACGGCTCTG ICGICCTGAA GGAAGIICTG GCCTAATICG CGGCGAAGGG TGICGCGIC CGCTCGCCIC	<u>gerbracadr cacacatrac car</u> gaacgat cc <mark>cctcg</mark> ctg gatgccgtgg tcaatggact tggcacggac cataceteae ggtecgtegg acgaecggag	XmnI ▼ aagaagttca cgcacgggcg ttccggagta cgggagttgt gaacggccgc gacgaagtcg gtcgcggctc ggcgggcggt gacgagcgag gtccggagga	agggttgtcg tacggagcgc actcaacgag gctccaggag	ggaggggttg aacccgccgc cgactggcct tcgccgcccg cgcggccgga gtatgtcatg tcgggggtga aatcaagcca ttcccccggg atcggctgtt
60 70 80 90 1234567890 1234567890 1234567890	CGGCGAAGGG	catacctcac	99c999c99t	tacggagcgc	aatcaagcca
70 1234567890	GCCTAATTCG	tggcacggac	gtcgcggctc		tcgggggtga
	GGAAGTUCTIG	tcaatggact	gacgaagtcg	gggcatacgg	. gtatgtcatg
40 567890 1234567890	TCGTCCTGBA	gatgccgtgg	gaacggccgc	. gtgacattgg	e66oo6oooo l
40 1234567890	CACGGCTCTG	cocctcoctg	cgggagttgt	gtgcgacgga	tegeegeeeg
30 1234567890	CGAGATTGCG	CATgaacgat	ttccggagta	cccaagtgag	: cgactggcct
10 20 30 1234567890 1234567890 1234567890 1234	ACATGCGTGA	CGCGCATTGC	cgcacgggcg	acgcgacgaa gcagccgaac cccaagtgag gtgcgacgga gtgacattgg	aacccgccgc
1234567890	GTCCCCCTCG	GCTATGGAGT	XmnI	acgcgacgaa	ggaggggttg

pikPKS Sequence

1800

1700

1600

pikPKS Sequence

1900

2000

REPLACEMENT SHEET

40 50 50 60 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 aacccttggt atagagcggg agacgacggg acaccatgga gaccacgcac accacgagcg ccacccctcccacgacacc catggcctgg ggtacacgcc aggtataggg ggaacgtagg gggaacgtagg gggaacatag gggggggt		Ctggggttgg gtgaaagcgc ggcttccgga gargtcrrch GccGGAATTA CCAGGACCGG TGCGAGAACA CCGGTGACAG GGCGTGGGGC M S S A G I T R T G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G R G A R T P V T G B R G R G A R T P V T G B A R T P V T G B A R T P V T G B A R T P T G B A R T P T G B A R T P T G B A R T R G T R G T V S R S E S E R V S V S N D A G S A H G T T G T G T G T G T G T G T G T G T	CCCGGGAC P V T CTCGTGCT R A R A CGCCGGTT A CGCCGGTT	TGCGAGAACA A R T CACTCCTTCT H S F S H S F S TGTCGAACGA	TA CCAGGACCGG T R T G GA TCATGCGGAG D H A E ECORI Apol MA TTCGTGTCCG	GCCGGAATTA A G I T CCGGCCCCGA G P D G P D EC AP GTCCGAGGAA	GAIGICITICA M S S TIGCCCCTG L P P A CCAAGAGIGA K S E	gacggagccg CCGACGGGGG RRG RRG RRG RTCGACAGTGT STVS	ggettec <u>gga</u> AaGTGCGGGT V R V TGGAGAGATG	gtgaaagcgc GaCaCGGGGG D T G E XmmI AATTGATTCG L I R	ctegg w w w GCCG
20 30 40 50 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 50 1234567890 567890 567890 1234567890 56780 5678		ag ggggggtgcc	gggagcat	ggaacgtagg	aggtataggg	ggtacacgcc	catggcctgg	tcccgacacc	ccggatagaa	ggggggt	&
40 50 50 70 80 70 80 70 80 50 70 80 50 50 1234567890 1234567890 1234567890 1234567890 1234567	• •	cg ccacccccg	accacgag	gaccacgcac	acaccatgga	agacgacgcg	atagagcggg	aacccttggt	tggatttccc	sctggcg	tta
		90 1234567890	123456789	80 1234567890	70 1234567890	60 1234567890	50 1234567890	40 1234567890	30 1234567890	20	1234

45/1	64
------	----

5	20	30	40	50	1234567890	70 1234567890	80 1234567890	10 20 30 40 50 60 70 80 90 100 100 100 100 100 100 100 100 100	100 1234567890	
<u>.</u>										
ပ္ပို့ ည	CGTCGCCGTC V A V	GCGGAACC CGTCGCCGTC GTCGGCATCT A E P V A V V G I S	CCTGCCGGGT C R V	GCCCGGCGCC P G A	cgggacccga r d p r	GAGAGTTCTG E F W	GGAACTCCTG E L L	CAGCGGAACC CGTCGCCGTC GTCGGCATCT CCTGCCGGGGT GCCCGGCGCCC CGGGACCTCTG GGAACTCCTG GCGCCAGGCG GCCAGGCCGT A E P V A V V G I S C R V P G A R D P R E F W E L L A A G G Q A V	GCCAGGCCGT Q A V	2100
GTC V	PACCGACGTC CCCGCGGACC G T D V P A D R	GCTGGAACGC W N A	CGGCGACTTC G D F	TACGACCCGG Y D P D	ACCGCTCCGC R S A	CCCCGGCCGC	TCGAACAGCC S N S R	CACCGACGTC CCCGCGGACC CGGCGACTTC TACGACCCGG ACCGCTCCGC CCCCGGCCGC TCGAACAGCC GGTGGGGCGG GTTCATCGAG 2200 T D V P A D R W N A G D F Y D P D R S A P G R S N S R W G G F I E	GTTCATCGAG F I E	2200
ACC R	GGTTCGACGC F D A	CGCCTTCTTC A F F	GGCATCTCGC G I S P	CCCGCGAGGC R E A	CGCGGAGATG A E M	GACCCGCAGC D P Q Q	AGCGGCTCGC R L A	GACGICGACC GGITCGACGC CGCCCGCGAGGC CGCGGAGATG GACCCGCAGC AGCGGCTCGC CCTGGAGCTG GGCTGGGAGG D V D R F D A A F F G I S P R E A A E M D P Q Q R L A L E L G W E A	GGCTGGGAGG G W E A	2300
ව්රි ස	CGCCGGGAIC A G I	CTGGAGCG CÉCCGGAIC GACCCGICCI L E R A G I D P S S	CGCTCACCGG L T G	CACCCGCACC T R T	GGCGTCTTCG G V F A	CCGGCGCCAT G A I	CTGGGACGAC W D D	CCCTGGAGCG CGCCGGGATC GACCCGTCACCGG CACCCGCACC GGCGTCTTCG CCGGCGCCAT CTGGGACGAC TACGCCACCC TGAAGCACCG 2400 L B R A G I D P S S L I G I R I G V F A G A I W D D Y A I L K H R	TGAAGCACCG K H R	2400
ည ဗ	CAGGGCGCCCCCCATCA Q G G A A I I	CCCCGCACAC P H T	CGTCACCGGC V T G	CTCCACCGCG L H R G	GCATCATCGC I I A	GAACCGACTC N R L	TCGTACACGC S Y T L	CCAGGGCGC GCCGCCATCA CCCCGCACAC CGTCACCGCG GCATCATCGC GAACCGACTC TCGTACACGC TCGGGCTCCG CGGCCCCAAGC Q G G A A I T P H T V T G L H R G I I A N R L S Y T L G L R G P S	CGGCCCCAGC G P S	2500

-1G. 31E

46/1	64
------	----

10	20	30	40	50	09	70	80	06	100	
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	
ATGGTCGTCG M V V D	3GICGICG ACICCGGCCA GICCTCGIC V V D S G Q S S S	GICCICGICG S S S	CTCGTCGCCG L V A V	TCCACCTCGC H L A	GTGCGAGAGC C E S	ATGGICGICG ACTCCGGCCA GICCTCGICGCG TCCACCTCGC GIGCGAGGC CTGCGGCGCG GCGAGICCGA GCICGCCCTC GCCGGGGGG 2600 M v v d s g o s s s l v a v h l a c e s l r r g e s e l a l a G v	GCGAGICCGA GCICGCCCTC GCCGGCGGCG ESELALAGGV	GCTCGCCTC L A L	GCCGGCGGCG A G G V	2600
TCTCGCTCAA S L N	TCGCTCAA CCTGGTGCCG GACAGCATCA S L N L V P D S I I	GACAGCATCA D S I I	TCGGGGCGAG G A S	CAAGTTCGGC K F G	GGCCTCTCCC	TCTCGCTCAA CCTGGTGCCG GACATCA TCGGGGCGAG CAAGTTCGGC GGCCTCTCCC CCGACGGCCG CGCCTACACC TTCGACGCGC GCGCCAACGG	CGCCTACACC A Y T	TTCGACGCGC F D A R	GCGCCAACGG A N G	2700
SnaBI CTACGTACGC Y V R	SnaBI ▼ TACGTACGC GGCGAGGGCG GCGGTTTCGT Y V R G E G G G F V	GCGGTTTCGT G F V	CGTCCTGAAG V L K	CGCCTCTCCC R L S R	GGGCCGTCGC A V A	SnaBI ▼ CTACGTACG GGCGAGGGCG GCGGTTTCGT CGTCCTGAAG CGCCTCTCCC GGGCCGTCGC CGACGGCGAC CCGGTGCTCG CCGTGATCCG GGGCAGCGCC 2800 Y V R G E G G F V V L K R L S R A V A D G D P V L A V I R G S A	CCGGTGCTCG P V L A	CCGTGATCCG V I R	GGGCAGCGCC G S A	2800
GTCAACAACG G	G A A	CGGCCCCC CCAGGGCATG	ACGACCCCG T T P D	ACGCGCAGGC A Q A	GCAGGAGGCC Q E A	GTCAACAACG GCGGCGCCGC CCAGGGCATG ACGCGCCAGGCG GCAGGAGGCC GTGCTCCGCG AGGCCCACGA GCGGGCCGGG ACCGCGCCGG	AGGCCCACGA A H E	GCGGGCCGGG R A G	ACCGCGCCGG T A P A	2900
CCGACGTGCG D V R	CCGACGTGCG GTACGTCGAG CTGCACGGCACCCCGATCGAGG CCGCTGCGCT CGGCGCCC CTCGGCACCG GCCGCCCGGC 3000 DVR YVE LHGT GT PVGDPIEA AALGAAL GFGRPA	CTGCACGGCA L H G T	CCGGCACCCC	CGTGGGCGAC V G D	CCGATCGAGG P I E A	CCGCTGCGCT A A L	CGGCGCCGCC	CTCGGCACCG L G T G	GCCGCCCGCC	3000

FIG. 31F

Sequence
pikPKS

		32	47/164	- .	ł
	3100	3200	3300	3400	3500
100 1234567890	GGCGGTCCGC A V R	GGTCGCGCGC TGCCCGCCAG CCTGAACTAC GAGACCCCGA ACCCGGCGAT CCCGTTCGAG GAACTGAACC TCCGGGTGAA CACGGAGTAC CTGCCGTGGG 3200		CGCGCTGGAC A L D	GGTGCTGTCG
90	AGGCCGTCCT	CACGGAGTAC CTGCCGTGGG	GAAGAGGCCC	AGTCCGCTGC	TGTCGATGCG
1234567890	A V L	T E Y L P W E	E E A P	S A A	V D A
80	CGGCATCGCC GGCCTCATCA AGGCCGTCCT	TCCGGGTGAA	TGTCGTGCTC	GTGTCGGCGA AGTCCGCTGC	ATGCGGGTGC
1234567890	G I A G L I K A V L	R V N	V V L	V S A K S A A	A G A
70	CGCCATCGCC	GAACTGAACC TCCGGGTGAA	CGAACGCGCA	GCCGTGGGTG	TGTCGACGCG GGCGCTGTCG ATGCGGGTGC TGTCGATGCG GGTGCTGTCG
	G I A	E L N L R V N	N A H	P W V	V D A G A V D A G A V D A G A V A
60	AGGGCGCGGC	CCCGTTCGAG	ATGGGCGGCA	GCGGTGTGGT	TGTCGACGCG
1234567890	G A A	P F E	M G G T	G V V	V D A
50	GCCACCTGG AGGGCGCGC	ACCCGGCGAT	CTCGTTCGGC	GCGGTCGGCG	NTC GTACGGATGG TGTCGACGCG GGCGCTGTCG ATGCGGGTGC TGTCGATGCG GGTGCTGTCG
1234567890		P A I	S F G	A V G G	R T D G V D A G A V D A G A V D A G A V A
40	GACGAACATC	GAGACCCCGA	TCGGCGTGTC	CGGCGGGTCG	TCGCGGGAIC
1234567890	T N I	E T P N	G V S	G G S	S R D R
30	GCTCGGTCAA	CCTGAACTAC GAG	CGGATGGTCG	AGTCGACGGT	CGCGTTCGCC TCC
1234567890	S V K	L N Y E	R M V V	S T V	A F A S
10 20 30 40 50 60 70 80 90 100 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	OGGACAGCCG CTCCTGGTCAA GACGAACATC GGCCACTGG AGGGCGCGGC CGGCATCGCC GGCCTCATCA AGGCCGTCCT GGCGGTCGC	TGCCCGCCAG P A S	AGCCGGAGCA CGACGGGCATGGTCG TCGGCGTGTC CTCGTTCGGC ATGGGCGGCA CGAACGCGCA TGTCGTGCTC GAAGAGGCCC CCGGGGGTTG PEHDGQRV VGVSSFGRGGCGTTCGGC TCGTTCGGC ATGGGCGGCA CGAACGCGCTC GAAGAGGCCC CCGGGGGTTG	TCGAGGTGCT TCGGTCGTGG AGTCGACGGT CGGCGGTCGCGG GCGGTGTGGT GCCGTGGGTG GTGTCGGCGA AGTCCGCTGC CGCGCTGGAC R G A S V V E S T V G G S A V G G G V V P W V S A K S A A A L D	GCGCAGATCG AGCGGCTTGC TCGCGGAAA Q I E R L A A F A S R D
1234567890	CGGACAGCCG	GGTCGCGCGC TGCCCGCCAGG G R A L P A S	AGCCGGAGCA P E H	TCGAGGTGCT R G A	GCGCAGATCG AGCGGCTTGC A Q I E R L A

FIG. 31G

48	/1	64
	, .	•

	3600	3700	000 88 80 80 80	I 3900	4000
1234567890	Bsu361 ▼ CGCCTGAGGG	GCTGGACTCT L D S	CCCGGTGCGC	CCCAGGCGGI Q A V	CAAGTCCAT
90	Eco47III GGGCTGGCCG A L A A	GTGCCGAACT A E L	ACGGCAGGCC R Q A	веветелсес в V т Р	CCCTGCGCAG
80 1234567890	TCTGGCGGCA L A A	GCCGGCATGG A G M G	AGGCCGTCGT A V V	GCAGCACCAC Q H H	CGTGTCGTGA
40 50 60 70 80 90 100 567890 1234567890 1234567890 1234567890 1234567890	ECO47III FGCCGTCGTC GTCGGCAGCG GGCCGGACGA TCTGGCGGCA GCGCTGGCCG A V V V G S G P D D L A A A L A A	TCTGGTCCGG GGCGTGGCTTCGGGG TTCGTGTTCC CCGGGCAGGG CACGCAGTGG GCCGGCATGG GTGCCGAACT GCTGGACTCT 3700 L V R G V A S G V G R V A F V F P G Q G T Q W A G M G A E L L D S	GTACGTCGAC TGGTCGCTGG AGGCCGTCGT Y V D W S L B A V V	CCACGCTGGA GCGGGTCGAT GTCGTGACGTT CGCCGTCATG GTCTCGCTGG CTCGCGTGTG GCAGCACCAC GGGGTGACGÇ CCCAGGCGGT Tlenne volvo v	CGTCGGCCAC TCGCAGGGCG AGATCGCCGC CGCGTACGTC GCCGGTGCCC TGAGCCTGGA CGACGCCGCT CGTGTCGTGA CCCTGCGCAG CAAGTCCATC
60 1234567890	GTCGGCAGCG V G S G	5 Q 5	GTACGTCGAC Y V D	GTCTCGCTGG V S L A	TGAGCCTGGA
50 1234567890	GGCCGTCGTC A V V	TTCGTGTTCC F V F P	CACTCTCCCC L S P	CGCCGTCATG A V M	GCCGGTGCCC
40 1234567890	TCGAGCACCG E H R	GCGAGTGGCG R V A	Bemi TGCGAGGCCG C E A A	CTGTGACGTT V T F	CGCGTACGTC
30 1234567890	CGTGCTCAGT R A Q F	CCGGTGTCGG G V G	CATGGCCGAA M A E	GTCGTGCAGC V V Q P	TCGCAGGGCG AGATCGCCGC
10 20 30 1234567890 1234567890 1234567890 1234	GGCCGGCGGG	CTGGTCCGG GGCGTGGCTT CCGGTGTCGG	TCGCGGCGGC A A A	ACGCTGGA GCGGGTCGAT GTCGTGCAGC T L E R V D V V Q P	TCGCAGGGCG
10 1234567890	ECO47III BBU361 T T R V L A G G R A Q F E H R A V V V G S G P D D L A A L A R P E G	rcregreege 1	BSMI TCCGCGGTGT TCGCGGCGGC CATGGCCGAGGCCG CACTCTCCCC GTACGTCGAC TGGTCGCTGG AGGCCGTCGT ACGGCAGGCC CCCGGTGCGC S A V F A A M A E C E A A L S P Y V D W S L E A V V R Q A P G A P	CCACGCTGGA T L E	CGTCGGCCAC

FIG. 31H

49	/1	64
----	----	----

100 567890	CGTCGCCG 4100 V A A	GGTCAT 4200 V I	GTGCCG 4300 V P	CGGCCG 4400	
90 100 1234567890 1234567890	EBP1 Sphi Bpull021 GCCGCCCACC TCGCCGGCAA GGGCGGCATG TCCGTCGCCGG A A H L A G K G G M L S L A L S E D A V L E R L A G F D G L S V A A	CTGTGAACGG GCCCACCGCC ACCGTGGTCT CGTACAGATC GAAGAGCTTG CTCGGGCGTG TGAGGCCGAT GGGGTCCGTG CGCGGGTCAT VNG PTATVVS GDPVQIE EELARRAC EADGVRA VI	ESPI BPU11021 T GGTCGAGATC ATCGAGAGCG AGCTCGCCGA GGTCCTCGCC GGGCTCAGCC CGCAGGCTCC GCGCGTGCCG VEILESELAEVLAEVLAPRVP	Kpni Acc65I ▼ ▼ ACCGAGCCCG TGCTCGACGG CGGCTACTGG TACCGCAACC TGCGCCATCG TGTGGGCTTC GCCCCGGCCG T E P V L D G G Y W Y R N L R H R V G F A P A V	
80 1234567890 1	TGGCCGGGTT (TGAGGCCGAT (Espi Bpull02I T GGGCTCAGCC CO	TGCGCCATCG R	
60 70 80 1234567890 1234567890	CTGGAGCGAC	CTCGGGCGTG R A C	GGTCCTCGCC V L A	KpnI Acc65I W W Y KG TACCGCAACC	
	I GGACGCCGTC D A V	GAAGAGCTTG E E L A	AGCTCGCCGA L A E	A CGGCTACTGG	
40 50 234567890 1234567890	Espi Bpull02I T CGCTGAGCGA (CGTACAGATC V Q I	ATCGAGAGCG.	TGCTCGACGG L D G	
~1	sphi ▼ crerccrcs r s r a	CCGGTGACCC G D P	GGTCGAGATC V E I		
30 1234567890	GGGGGCATG G G M	ACCGTGGTCT T V V S	ACAGCCGGCA S R Q	CGCCTGGATC A W I	
10 20 30 1234567890 1234567890 1234567890	TCGCCGGCAA A G K	GCCCACCGCC P T A	Mlui TCCCGTCGAC TACGCGTCCC ACAGCCGGCA P V D Y A S H S R Q	TTCTTCTCGA CACTCGAAGG CGCCTGGATC	MacI
10 1234567890	GCGCCCACC	CTGTGAACGG V N G	TCCGTCGAC P V D	TICTICICGA F F S T	

FIG. 311

5000

GCTGCTCCTC GTCGTGCACG

10	20	30	40	50	09	70	80	06	100	
1234567890	12345678	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
GACCCTGCGT T L R	MaeI BfaI ▼ GACCCTGCGT CGCGACAACG GCGCTAGGTC GCCTCCCTCG CCGAAGCATG GGCCAACGGA CTCGCGGTCG ACTGGAGCCC GCTCCTCCC 4600 T L R R D N G G Q D R L V A S L A E A W A N G L A V D W S P L L P	GCGGTCAGGA G Q D	Mael Bfal T CCGCCTAGTC R L V	GCTCCCTCG A S L A	Maei Bfai ▼ CTAGTC GCCTCCTCG CCGAAGCATG	GGCCAACGGA A N G	CTCGCGGTCG	GGCCAACGGA CTCGCGGTCG ACTGGAGCCC A N G L A V D W S P	GCTCCTCCCC L L P	4600
TCCGCGACCG S A T G	MIUI TCCGCGACCG GCCACCACTCCCC ACCTACGCGT TCCAGACCGA GCGCCACTGG CTGCGCGAGCGAGT TCGAGGCGCT CGCCCCGGCG GGCGAGCCGG 4700 S A T G H H S D L P T Y A F Q T E R H W L G E I E A L A P A G E P A	CGACCTCCCC D L P	MluI A ACCTACGCGT T Y A F	TCCAGACCGA Q T E	GCGCCACTGG R H W	CTGGGCGAGA L G E I	TCGAGGCGCT E A L	CGCCCCGGCG A P A	GECGAGCCGG	4700
CGGTGCAGCC V Q P	CGGTGCAGCC CGCCGTCCTC CGCACGAGGC GGCGGAGCTC GACCGGGACG AGCAGCTGCG CGTGATCCTG GACAAGGTCC GGGCGCAGAC	CGCACGGAGG R T E A	CGGCCGAGCC A B P	GGCGGAGCTC A E L	GACCGGGACG AGCAGCTGCG D R D B Q L R	AGCAGCTGCG Q L R	CGTGATCCTG V I L	CGTGATCCTG GACAAGGTCC GO	GGGCGCAGAC A Q I	4800
GGCCCAGGTG A Q V	GGCCCAGGTG CTGGGGTACG CGACAGGATCGAG GTCGACCGGA CCTTCCGTGA GGCCGGTTGC ACCTCCCTGA CCGGCGTGGA CCTGCGCAAC A Q V L G Y A T G G Q I E V D R T F R E A G C T S L T G V D L R N	CGACAGGCGG	GCAGATCGAG Q I E	GTCGACCGGA V D R T	CCTTCCGTGA F R E	GGCCGGTTGC A G C	ACCTCCCTGA T S L T	BECCEGITGC ACCICCTGA CCGGCGIGGA CCTGCG	CCTGCGCAAC L R N	4900
							-		Anal.T	

pikPKS Sequence

FIG. 31J

CGGATCAACG CCGCCTTCGG CGTACGGATG GCGCCGTCCA TGATCTTCGA CTTCCCCACC CCCGAGGCTC TCGCGGAGCA R I R D F P T P E A L A E Q

5	1/	1	64
J	'/	•	U T

	5100	5200	5300	5400	9200
100	CTGCCGCCT	GGACGTGGAG	TTCTTCGGGA	CGACCTCCCT	CCTGCTGACC
	C R L	D V E	F F G I	T S L	L L T
90 1234567890	GTCGCCATGG V G M A	ACCGCGGCTG (R G W	CGACGCGGCC	CCACCCC GGCATCCACC CCACCTCC E D A G I D P T S	TCGACGGCTA D G Y
40 50 60 70 80 90 100	3 GCCGGTGCCG TCGACGAGCC GGTGGCGATC GTCGGCATGG CCTGCCGCCT	TTCCCGCAGG ACCGCGGCTG	TCGCCGGCTT	CGAGGACGCC	GGGGAAGGCC
567890 1234567890 1234567890 1234567890 1234567890 1234567890	A G A V D E P V A I V G M A C R L	F P Q D R G W	A G F	E D A	
70	TCGACGAGCC	GATCTCGGAG	ATCGAGAACG	GGGAGGCCGT	GCGGGACGGC
1234567890	D E P	I S E	I E N V	E A V	R D G
60	GCCGGTGCCG	SCGGCTG GTGGCCGCCG GCGCGACGC GATCTCGGAG	GGGCGGTTTC	GCTCCTCCTC GAAACCTCCT	GGCCGAGCCT
1234567890	A G A V	R L V A G G G D A I S E	G G F	L L L E T S W	P S L
50	GCTCC GGTGGCGGCG	GTGGCCGGCG	ACGTCCGCCA	GCTCCTCCTC	CACGAGTACG
1234567890	A P V A A	V A G G	V R Q	L L L	H E Y G
40	AGCCGGCTCC	GTGGCGGCTG	GGCACGTCGT	CGCAGCAGCG	GGCGATGACC
1234567890	P A P	WRL	G T S Y	Q Q R	A M T
30	GCCGGTGCGG	CGGAGGACCT	GGAGCACCCC	GCCATGGACC	TCTTCACTGG
	A G A E	E D L	E H P	A M D P	F T G
1234567890 1234567890 1234567890 1234	GGGAGGCGC GGCGAACCCG GCCGGTGCGG AGCCGCTGC GGTGGCGATC GTCGCCGATC GTCGCCGCTTGG CCTGCCGCCTT Baaany bagabaccagaccagaccagaccagaccagaccagacca	GCCCGGTGGG GTCGCCTCGC CGGAGGACCT GTGGCGGCTG GTGGCGGACGC GATCTCGGAG TTCCCGCAGG ACCGCGGCTG GGACGTGGAG 5200 P G G V A S P E D L W R L V A G G G D A I S E F P Q D R G W D V E	BAMHI GGGCTGTACC ACCCGGAGCACCCCC GGCACGTCCT ACGTCCGCCA GGGCGGTTTC ATCGAGAACG TCGCCGGCTT CGACGCGGCC TTCTTCGGGA G L Y H P D P E H P G T S Y V R Q G G F I E N V A G F D A A F F G I	TCTCGCCGCG CÁAGGCCCTC GCATCGAGCG GCTCCTCCTC GGAACCTCCT GGGAGGCCGT CGAGGACGCC GGGATCGACC CGACCTCCCT S P R E A L A M D P Q Q R L L L E T S W E A V E D A G I D P T S L	GCGGGGACGG CAGGTCGGCG TCTTCACTGG GGCGAGTACG GGCCGAGCCT GCGGGACGGC GGGGAAGGCC TCGACGGCTA CCTGCTGACC R G R Q V G V F T G A M T H E Y G P S L R D G G E G L D G Y L L T
1234567890	GGGAGGCGGC B A A	SCCCGGTGGG P B G	GGCTGTACC	TCTCGCCGCG S P R	GCGGGGACGG R G R

pikPKS Sequence

0009

GGACGTCGTC GAGGCACACG GCACGGGCAC
D V V E A H G T G T

52/1	64
9	

890	CG 5600	TT 5700 F	, 0088 OT.	TC 5900 P	
1 12345678	TCGCTGGT S L V	TCGTCGAG V B	ccrccrcs v 1 v	CTCACGGC L T A	
90 1234567890	CTGCTCGTCG C S S	CCCGGGATGT P G M F	GCGTCGGCGT V G V	GAGCAACGGC S N G	
80 1234567890	TGGACACGGC D T A	GATGCCCACG M P T	TGGTCCGAGG W S E G	AGGACGCCC D G A	
10 20 30 40 50 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	GGCAACACGG CCAGCGTCAT GTCGGCCG CACTCGGCCT TGAGGGCCCC GCCCTGACGG TGGACACGGC CTGCTCGTCG TCGCTGGTCG 5600	CCCTGCACCT CGCCGTGCAG GCCCTGCGCA AGGGCGAGGT CGACATGGCG CTCGCCGGCGG GCGTGGCCGT GATGCCCACG CCCGGGATGT TCGTCGAGTT L H L A V Q A L R K G E V D M A L A G G V A V M P T P G M F V B F	XMUI CAGCCGGCAG CGCGGGCACGG CCGGTCGAAG GCGTTCGCCG CGTCGGCGAAC CGGCACCAGC TGGTCCGAGG GCGTCGGCGT CCTCCTCGTC S R Q R G L A G D G R S K A F A A S A D G T S W S E G V G V L L V	GÁGCGCCTGT CGGACGCCCG CCGCAACGGA CACCAGGTCG TCGCGGCAGC GCCTTGAACC AGGACGGCGC GAGCAACGGC CTCACGGCTC E R L S D A R R N G H Q V L A V V R G S A L N Q D G A S N G L T A P	
60 1234567890	TGAGGGCCCC B G P	CTCGCCGGCG L A G G	CGTCGGCGGA S A D	CCGCGGCAGC R G S	
50 1234567890	CACTCGGCCT L G L	CGACATGGCG D M A	XmnI GCGTTCGCCG A F A A	TCGCGGTCGT A V V	
40 1234567890	GTCTCGTACA V S Y T	AGGGCGAGGT G E V	CCGGTCGAAG R S K	CACCAGGTCC H Q V L	
30 1234567890	GTCGGGCCGC S G R	COCCOTICAG GCCTGCGCA AGGG A V Q A L R K G	S C S	CCGCAACGGA CACC R N G H Q	
20 1234567890	GGCAACACGG CCAGCGTGAT (CGCCGTGCAG A V Q	AGCCGGCAG CGCGGGACGG S R Q R G L A G D G	CGGACGCCCG D A R	
10 1234567890	GCCACACGG G N T A	CCCTGCACCT C	CAGCCGGCAG S R Q	GÁGCGCCTGT (E R L S	

pikPKS Sequence

FIG. 31L

CGAACGGGCC CTCGCAGCAG CGCGTCATCC GGCGCGCGT GGCGGACGCC CGGCTGACGA CCTCCGACGT N G P S Q Q R V I R R A L A D A R L T T S D V

Ü
⊑
Ф
3
5
Φ
Ś
S
Ţ
О.,
×
-
Q

		JEININE II	53/164	•	
	6100	6200	6300	6400	6500
100 1234567890		GTCGACGAGC V D E P	GGGCCGCCGT A A V	CICCICCTIC GGGAICAGCG GCACCAAIGIGGIG CICGAAGAGG CCCCGGIGGI IGICGAGGGI GCITCGGICG ICGAGCCGIC GGITGGCGGG S S F G I S G I N A H V V L E E A P V V E G A S V V E P S V G G	Bemi TCGCCGGTCG GCGCCGCTTGG GTGGTGTCGG CGAAGTCCGC TGCCGCGCTC GACGCGCAGA TCGAGCGGCT TGCCGCATTC GCCTCGCGGG S A V G G G V T P W V V S A K S A A L D A Q I E R L A A F A S R D
90 1234567890	TCGGGTCGTT G S L	GACGCTGCAC T L H	GGCTGCGCC	GCTTCGGTCG TCGAGCCGTC A S V V E P S	BSMI TCGAGCGGCT TGCCGCATTC GCCTCG ERLARS
80	CCGCTGCGCC	TGCTGCCGAA	GCAGGACGGC	GCTTCGGTCG	TCGAGCGGCT
1234567890		L P K	Q D G	A S V V	E R L
70	CGACGAACAG	CCAGGCGATG CGCCACGGAC TGCTGCCGAA	GGCCGGAGAA	TGTCGAGGGT	GACGCGCAGA
1234567890	D E Q	Q A M R H G L L P K	P E K	V E G	D A Q I
60	AGGCCGTGA	CCAGGCGATG	GCCGTCGACT	CCCCGGTGGT	TGCCGCGCTC
1234567890	G R D	Q A M	A V D W	P V V	A A L
50	ACCTACGGCC AGGGCCGTGA CGACGAACAG CCGCTGCGCC TCGGGTCGTT	TCAAGATGGT	CCTCACCGAG	CTCGAAGAGG	CGAAGTCCGC
123 4 567890	T Y G Q G R D D E Q P L R L G S L	K M V	L T E	L E E A	K S A
40	CCTGATCGCC	TCCGGTGTCA	CCGTGGAACT	GCATGTGGTG	GTGGTGTCGG
1234567890	L I A	S G V I	V E L	H V V	V V S A
30	AGGCGCAGGC	GGCCGCGTC A G V	TCGGCTGGCG S A G A	GCACCAATGC T N A	GACGCCTTGG T P W
10 20 30 40 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	REACTICGGC GACCCGATCG AGGCGCAGGC	CCCAGGCCGC Q A A	CCTCGGACCA GATCGACTGG TCGGCTGGGG CCGTCGAGG GCCGTCGACT GGCCGGAGAA GCAGGACGGC GGGCTGCGCC GGGCCGCCGT S D Q I D W S A G A V E L L T E A V D W P E K Q D G G L R R A A V	TCCTCCTTC GGGATCAGCG GCACCAATGC S S F G I S G T N A	GCGGTCG GCGGCGTGT A V G G G V
1234567890	GCGACTCGGC	ATCGGGCACA	CCTCGGACCA	CICCICCITIC	TCGGCGGTCG
	R L G	I G H T	S D Q	S S F	S A V G

FIG. 31M

54	1 (54	+

			•		
	0099	6700	0089	0069	7000
1234567890	ccgrcecgcr V A L	Xmni * *GCGCCCGGG GCGGACGACC TCGTACAGGC GCTGGCCGAT CCGGACGGC TGATACGCGG AACGGCTTCC GGTGTCGGGC GAGTGGCGTT CGTGTTCCCC 6700 G A G A D D L V Q A L A D P D G L I R G T A S G V G R V A F V F P	CTGTCCCGT L S P Y	CCGTCATGGT V M V	CTCGCTGGCT CGCGTGTGGC AGCACCACGCCC CAGGCGGTCG TCGGCCACTC GCAGGGCGAG ATCGCCGCCG CGTACGTCGC CGGAGCCCTG S L A R V W Q H H G V T P Q A V V G H S Q G E I A A A Y V A G A L
90	CACGTACTGG CTGACGGGGG TGCTCAGTTC GAGCACCGGG CCGTCGCGCT	GAGTGGCGTT	TGAGGCCGCG	GTGACGTTCG	CGTACGTCGC
1234567890		V A F	E A A	V T F A	Y V A
80	TGCTCAGTTC	GGTGTCGGGC	TGGCCGAGTG	CGTGCAGCCT	ATCGCCGCCG
1234567890	A Q F	GVGR	A E C	V Q P	I A A A
70	CTGACGGGCG	XmnI ▼ AACGGCTTCC T A S	GCGGCGGCCA	GGGTCGATGT	GCAGGGCGAG
1234567890	D G R		A A A M	V D V	Q G E
60	CACGTACTGG	TGATACGCGG	CGCGGTGTTC	ACGCTGGAGC	TCGGCCACTC
1234567890	H V L A	I R G	A V F	T L E R	G H S
50	CGCGGG CGCTGTCGCT	CCGGACGGGC P D G L	TGGACTCTTC D S S	CGGTGCGCCC G A P	CAGGCGGTCG Q A V V
40	TCGACGCGGG	GCTGGCCGAT	GCCGAACTGC	GGCAGGCCC	TGTGACGCCC
	D A G	L A D	A E L L	Q A P	V T P
30	GCCGGTGCTG	TCGTACAGGC	TGGCATGGGT	GCCGTCGTAC	AGCACCACGG
	A G A V	V Q A	G M G	A V V R	H H G
10 20 30 60 70 80 90 100 100 20 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890	ATCETACEGA TGACGCCGAC GCCGGTGCTG CGCTGTCGCT CACGTACTGG CTGACGGGCG TGCTCAGTTC GAGCACCGGG CCGTCGCCT	GCGGACGACC A D D L	SGTCAGGGCA CGCAGTGGCATGGGT GCCGAACTGC TGGACTCTTC CGCGGTGTTC GCGGCGGCCA TGGCCGAGTG TGAGGCCGCG CTGTCCCCGT 3 Q G T Q W A G M G A E L L D S S A V F A A A M A E C E A A L S P Y	ACGICGACTG GICGITGIAC GGCAGGCCCC CGGTGCGCC ACGCTGGAGC GGGTCGATGT CGTGCAGCT GTGACGTTCG CCGTCATGGT V D W S L E A V V R Q A P G A P T L E R V D V V Q P V T F A V M V	CGCGTGTGGC R V W Q
10	ATCGTACGGA R T D	3GCGCCGGG	3GTCAGGGCA 3 Q G T	ACGICGACTG V D W	CTCGCTGGCT S L A

55/1	164
------	-----

10	20	30	40	20	09	70	08	90	100	
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1	1234567890	1234567890	123456/890	123456/890	
CCCCTGGACG P L D D	ACGCCGCCG A A R	CGTCGTCACC V V T	CTGCGCAGCA L R S K	AGTCCATCGC S I A	Sphi CCCCTGGACG ACGCCGCCC CGTCGTCACC AGTCCATCGC CGCCCACCTC GCCGGCAAGG GCGGCATGCT GTCCCTCGCG CTGAACGAGG 7100 P L D D A A R V V T L R S K S I A A H L A G K G G M L S L A L N E D	GCCGGCAAGG A G K G	Sphi GCGGCATGCT G' G M L	I GTCCCTCGCG S L A	CTGAACGAGG L N E D	7100
ACGCCGTCCT A V L	GGAGCGACTG E R L	AGTGACTTCG S D F D	ACGGGCTGTC G L S	CGTCGCCGCC V A A	ACGCCGICCT GGAGCGACTG AGGGGCTGTC CGTCGCCGCC GICAACGGGC CCACCGCCAC TGTCGTGTCG GGTGACCCCG TACAGATCGA 7200 A V L E R L S D F D G L S V A A V N G P T A T V V S G D P V Q I E	CCACCGCCAC T A T	TGTCGTGTCG V V S	GGTGACCCCG G D P V	TACAGATCGA Q I E	7200
AGAGCTTGCT E L A	GAGCTTGCT CAGGCGTGCA AGGCGGACGG E L A Q A C K A D G	AGGCGGACGG A D G	ATTCCGCGCG F R A	CGGATCAITC R I I P	Mlui $lacktright$ AGAGCGTGCA AGGCGGACGG ATTCCGCGCG CGGATCATTC CCGTCGACTA CGCGTCCCAC AGCCGGCAGG TCGAGATCAT CGAGAGCGAGG E LA QACK ADG FRARILIPVDYAS B	MluI CGCGTCCCAC A S H	AGCCGGCAGG S R Q V	TCGAGATCAT E I I	CGAGAGCGAG ESE	7300
CTCGCCCAGG L A Q V	CTCGCCCAGG TCCTCGCCGG	TCTCAGCCCG L S P	CAGGCCCCGC Q A P R	GCGTGCCGTT V P F	CTCGCCCAGG TCCTCGCCGC CAGGCCCCGC GCGTGCCGTT CTTCTCGACG CTCGAAGGCA CCTGGATCAC CGAGCCCGTC CTCGACGGCA 7400 L A Q V L A G L S P Q A P R V P F F S T L E G T W I T E P V L D G T	CTCGAAGGCA L E G T	CCTGGATCAC W I T	CGAGCCCGTC E P V	CTCGACGGCA L D G T	7400
Acc ▼ CCTACTGGTA Y W Y	KpnI Acc651 W GTA CCGCAACCTC Y R N L	CGTCACCGCG R H R V	TCGGCTTCGC G F A	CCCCGCCATC P A I	KpnI Acc65I ▼ ▼ CCTACTGGTA CCGCAACCTC CGTCGCCGTCGCCATC GAGACCCTGG CCGTCGACGA GGGCTTCACG CACTTCGTCG AGGTCAGCGC	CCGTCGACGA V D E	GGGCTTCACG	CACTICGICG H F V E	AGGTCAGCGC V S A	7500

pikPKS Sequence

8000

TCGCTCCTCG

CGGCGTGGTC

>

Ö

US	9/	900	5,30) 4			
_		_	_				

56/164				
047III	7700	7800		
90	Ft H	ي ا		

7600	Eco47III	7800	7900
CCACCCCGTC CTCACCATGA CCCTCCCCGA GACCGTCACC GGCCTCGGCA CGCAAGAGA GGCCAAGAGC GTCTGGTCAC CTCGCTCGCC 7600 H P V L T M T L P B T V T G L G T L R R B Q G G Q B R L V T S L A	E A W V N G L P V A W T S L L P A T A S R P G L P T Y A F Q A E R Y	xhol PaeR7I ▼ ACTGGCTCGA GACTGGCGC TACCGCATCG ACTGGAAGCG CCTCCCGGCC GCCGAGGGGT CCGAGCGCAC 7800 WLENTPA LATGDDWRYRIDWKRLPA EGSERT	CGGCCTGTCC GGCCGCTGCC TCGCCGTCAC GCCGGAGGAC CACTCCGCGC AGGCCGCCGC CGTGCTCAC GCGCTGGTCG ACGCCGGGGC GAAGGTCGAG
GTCTGGTCAC	CGCCTTCCAG	GCCGAGGGGT	ACGCCGGGC
L V T	A F Q	A E G S	A G A
GGCCAAGAGC	TGCCCACCTA	CCTCCCGGCC	GCGCTGGTCG
G Q E R	P T Y	L P A	A L V D
CGAACAGGGA	CGCCCCGGTC	ACTGGAAGCG	CGTGCTCACC
B Q G	R P G L	W K R	V L T
CCCTCCGTCG	CACGGCCTCC	TACCGCATCG	AGGCCGCCGC
L R R	T A S	Y R I D	A A A
GCCTCGGCA G L G T	TCCTGCCGC L P A	ali Paccegoga ceactegage taccecatog actegaages cotaccego secogages cogages. T g d d w r y r i d w k r l p a a e g s e r t	CACTCCGCGC H S A Q
GACCGTCACC	TGGACTTCGC	MBCI Bali CCACCGGCGA T G D	GCCGGAGGAC
T V T	W T S L		P E D
CCCTCCCCGA	TCCCGTGGCA	GCCGCCTGG	TCGCCGTCAC
L P E	P V A	A A L A	A V T
CACCCCGTC CTCACCATGA CCCTCCCCGA GACCGTCACC GGCCTCGGCA CCCTCCGTCG CGAACAGGGA GGCCAAGAGC GTCTGGTCAC CTCGCTCGC H P V L T M T L P E T V T G L G T L R R E Q G G Q E R L V T S L A	TCAACGGGCT N G L	XhoI PaeR7I ▼ TGGCTCGA GAACACTCCC GCCGCCTGG W L E N T P A A L A	GGCCGCTGGC GRWL
CCACCCGTC H P V	GAGGCGTGGG E A W V	xhol PaeR7I ▼ ACTGGCTCGA G	CGGCCTGTCC

pikPKS Sequence

GAGGCCCTCG CCGCCCGGCT CACCGCCTC ACGCCTTCAC E A L A L T G D G F T

GTGCTGACGG CCGGGGGGA CGACGACCGT

×
\simeq
<u>-</u>
9
3
℧
0
ဟ
S
×
₽.
¥
=
•

	8100	8200	57/164 ຂໍ້	8400	8500
100					
1234567	ccercee.	CGTCGACC V D L	GGACTCC G L H	GAGCCCT A L	ACTCACC L T
90	GGCGCGGTCT	GGGCCGGCCT	CCGCACCACC	GGCGGCACCG	GAGCCACCCA
1234567890	G A V S	A G L	R T T	G G T G	A T O
10 20 30 40 50 60 70 80 90 100	ACGGACTOGT ACCGCAGGTC GCCTGGGCTCGG CGACGCCGGA ATCAAGGCGC CCCTGTGGTC CGTCACCCAG GGCGCGGTCT CCGTCGGACG	TCTCGACACC CCGCCGACCGGGC CATGCTCTGG GGCCTCGGCC GCGTCGTCGC CCTTGAGCAC CCCGAACGCT GGGCCGGCCT CGTCGACCTC	BBABI CCCGCCCAGC CCCATCGCC CACCTCGTCA CCCCACTCTC CGGCGCCACC GGCGAGGACC AGATCGCCAT CCGCACCACC GGACTCCACG PAQPDAAL TALSGATCCACCACCACCACCACCACCACCACCACCACCACCACC	CCCGCCGCCT CGCCCGCCAC CCCTCCACG GACGTCGGCC CACCGCGAC TGGCAGCCCC ACGGCACCGT CCTCATCACC GGCGGCACCG GAGCCCTCGG	CAGCCACGCC GCACGCTGGA TGGCCCACCA CGGAGCCGAA CACCTCCTCC TCGTCAGCCG CAGCGGCGAA CAAGCCCCCC GAGCCACCCA ACTCACCGCC
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	G L V P Q V A W V Q A L G D A G I K A P L W S V T Q G A V S V G R	L D T P A D P D R A M L W G L G R V V A L E H P E R W A G L V D L		RRLARR GRRPTRD WQPHGTVLITGGTG	S H A A R W M A H H G A E H L L L V S R S G E Q A P G A T Q L T A
70	CCCTGTGGTC	CCTTGAGCAC	GGCGAGGACC	ACGGCACCGT	CAGCGGCGAA
	L W S	L E H	G E D Q	G T V	S G E
60	ATCAAGGCGC CCCTGTGGTC	GCGTCGTCGC	CGGCGCCACC	TGGCAGCCCC	TCGTCAGCCG
1234567890	I K A P L W S	V V A	G A T	W Q P H	V S R
50	CGACGCCGGA	GCTCGCC GCGTCGTCGC	CCGCACTCTC	CACCCGCGAC	CACCTCCTCC
1234567890	D A G		A L S	T R D	H L L L
40	AGGCGCTCGG	CATGCTCTGG	CACCTCGTCA	GACGTCGGCC	CGGAGCCGAA
	A L G	M L W	H L V T	R R P	G A E
30	GCCTGGGTCC A W V Q	CCGACCGGGC D R A	CGCCTCGCC A L A	CCCCTCCACG P L H G	TGGCCCACCA
20 1234567890	GENCE ACCGCAGGTC GCCTGGGTCC AGGCGCT	CCCGCCGACC CCGACCGGGC CATGCT	CCCGCCCAGC CCCCCTCGCC CACCTCG	CGCCCGCGCA A'RA	AGCCACGCC GCACGCTGGA TGGCCCACCA CGGAGC S H A A R W M A H H G A
10	ACGGACTCGT	TCTCGACACC	CCCGCCCAGC	CCCGCCGCCT	CAGCCACGCC
1234567890	G L V	L D T	P A Q P	R R L	

58/1	64
------	----

10	20	30	10 10 20 30 40 50 60 70 80 100 100 100 100 100 1234567890 123456780 123456780 123456780 123456780 123456780 123456780 123456780 123456780 123456780 123456780 123456780 123456780 123456780	50 1234567890	60 1234567890	70 1234567890	80 1234567890	90 1234567890	1234567890	
	していています。	CCGCGTCACC	ATCGCCGCT	GCGACGTCGC	CGACCCCAC	GCCATGCGCA	CCCTCCTCGA	CGCCATCCCC	GCCGAGACGC	8600
E L A A	S G A	E+ >	ы ч	D V A	н ф О	A A T	r r O	а н ч	ELTA SGARVTIAAC DVA DPHAMRT LLD AIPAETP	
CCTCACGC L T A	CGTCGTCCAC V V H	ACCGCCGGCG T A G A	CGCTCGACGA L D D	CGGCATCGTG G I V	GACACGCTGA D T L T	CCGCCGAGCA A E Q	GGTCCGGCGG V R R	GCCCACCGTG A H R A	CCCTCACCGC CGTCGTCGCC ACCGCCGGCG CGCTCGACGA CGGCTGAGCTGCGAGCA GGTCCGGCGG GCCCACCGTG CGAAGGCCGT L T A V V H T A G A L D D G I V D T L T A E Q V R R A H R A K A V Mlui	8700
CGGCGCCTCG G A S	GTGCTCGACG V L D E	AGCTGACCCG L T R	GGACCTCGAC D L D	CTCGACGCGT L D A F	TCGTGCTCTT V L F	CTCGTCCGTG S S V	TCGAGCACTC S S T L	TGGGCATCCC G I P	CGGCGCCTCG GTGCTCGACCCCG GGACCTCGACGCGT TCGTGCTCTT CTCGTCCGTG TCGAGCACTC TGGGCATCCC CGGTCAGGGC 8800 G A S V L D E L T R D L D A F V L F S S V S S T L G I P G Q G	8800
AACTACGCCC N Y A P	CGCACAACGC H N A	CTACCTCGAC Y L D	GCCCTCGCGG A L A A	CTCGCCGCCG R R R	GGCCACCGGC A T G	CGGTCCGCCG R S A V	TCTCGGTGGC S V A	CTGGGGACCG W G P	AACTACGCCC CGCACAACGC CTACCTCGCGG GCCCACCGGC CGGTCCGCCG TCTCGGTGGC CTGGGGACCG TGGACGTGG N Y A P H N A Y L D A L A A R R A T G R S A V S V A W G P W D G G	8900
GCGGCATGGC G M A	CGCCGGTGAC A G D	GGCGTGGCCG G V A B	GCGGCATGGC CGCCGGTGAC GGCGTGGCTGCG CAACCACGGC GTGCCCGGCA TGGACCCGGA ACTCGCCCTG GCCGCACTGG AGTCCGCGCT G M A A G D G V A E R L R N H G V P G M D P E L A L A A L E S A L	CAACCACGGC N H G	GTGCCCGGCA V P G M	TGGACCCGGA D P E	ACTCGCCCTG L A L	GCCGCACTGG A A L E	AGTCCGCGCT S A L	0006

FIG. 31R

59/1	64
00	

GGTCGCCCGC AGCCCCTC	υ	GTACTCCTC' Y S S	TCTACCTCGC GTACTCCTC Y L A Y S S	TGGGACCGCT TCTACCTCGC GTACTCCTC	GGACATCGAC TGGGACCGCT TCTACCTCGC GTACTCCTC D I D W D R F Y L A Y S S	TCACCGICGC GGACAICGAC IGGGACCGCT ICTACCICGC GIACICCIC	CGGCCGGGAC GAGACCGTCGC GGACATCGAC TGGGACCGCT TCTACCTCGC GTACTCCTCC GGTCGCCCGC AGCCCCTCGT CGAGGAGCTG G R D E T A I T V A D I D W D R F Y L A Y S S G R P Q P L V E E L
BetXI GCGCCAACCC CCTGGCCC A N P L A E	99 0	TCCGCCC S A Q	GGGCGGGAGC TCCGCCC	CGTCCGGACA GGGCGGGAGC TCCGCCC? S G Q G G S S A Q	GACAGCGCCA CGTCCGGACA GGGCGGGAGC TCCGCCCA D S A T S G Q G G S S A Q	CGACGCACGG GACAGCGCCA CGTCCGGAACA GGGCGGGAGC TCCGCCC? D A R D S A T S G Q G G S S A Q	BBLXI CCCGAGGTGC GGCGCATCAT CGACAGCGCCA CGTCCGGACA GGGCGGGAGC TCCGCCCAGG GCGCCAACCC CCTGGCCGAG CGGCTGGCCG 9200 PEVRRIID ARD SATS G Q G G S A Q G A N PLAER LAA
GATECGTTCG CCGGAGG	ပ္ပို့ 🕊	CCGTGCT V L	CAGGCGCCG CCGTGCT Q A A A V L	CGTACGGGCG CAGGCCGCCG CCGTGCT V R A Q A A A V L	TCCTCGGTCT CGTACGGGCG CAGGCCGCCG CCGTGCT	ACGGAGATCC TCCTCGGTCT CGTACGGGCG CAGGCCGCCG CCGTGCT TEILLGLVRAQAAAVL	CCGCGCCTCC CGGCGAGCGT ACGGAGATCC TCCTCGGTCT CGTACGGCG CAGGCCGCCG CCGTCCG GATGCGTTCG CCGGAGGACG TCGCCGCCGA 9300 A A P G E R T E I L L G L V R A Q A A A V L R M R S P E D V A A D
ACCGGGCTCC AGCTGCC(T G L Q L P	A A	GACCCGG T R	GCAACAGGCT GACCCGG N R L T R	GTCGAGCTGC GCAACAGGCT GACCCGG V E L R N R L T R	GCTCGCCGGT GTCGAGCTGC GCAACAGGCT GACCCGG	GCTTCGACTC GCTCGCCGGT GTCGAGCTGC GCAACAGGCT GACCCGG FDSLAGVELRNRLTR	CCGCGCCTTC AAGGACATCG GCTTCGACTGC GCGAGCTGC GCAACAGGCT GACCCGGGCG ACCGGGCTCC AGCTGCCCGC GACGCTCGTC R A F K D I G F D S L A G V E L R N R L T R A T G L Q L P A T L V
CGGACGCCG GCGGTCC	ACGG I A	GAGGAG	CCTCGGTGAC GAGGAG	GCAGCGAGTT CCTCGGTGAC GAGGAG	TCGCTGCTCC GCAGCGAGTT CCTCGGTGAC GAGGAG	GGCCCTCGTG TCGCTGCTCC GCAGCGAGTT CCTCGGTGAC GAGGAG	TICGACCACC CGACGCCGCT GCCTGCTCC GCAGCGAGTT CCTCGGTGAC GAGGAGGCGC CGGACGCCCG GCGGTCCGCG GCGCTGCCCG

FIG. 31S

REPL	ACEN	JENT	SHEET
------	------	-------------	-------

60/1	64

	0096	9700	9800	0066
100	GTGACATCCG	CAGCCCGGAG GACCTGTGC GGATGCTGTC CGAGGGCGGC GAGGGCATCA CGCCGTTCCC CACCGACCGC GGCTGGGACC TCGACGGCCT GTACGACGCC	GACCCGGACG CGCTCGGCA GGCGTC CGCGAGGGCG GGTTCCTGCA CGACGCGGCC GAGTTCGACG CGGCGTTCT CGGCGTTCT CGGCGTTCG CCGCGCGAGG	GCAGCAGCAC
1234567890	D I R		D P D A L G R A Y V R E G G F L H D A A B F D A E F F G V S P R E A	S S T
90	CGCTACCCCG	TCGACGGCCT	CGGCGTCTCG	TCGCTGCGCG
1234567890	R Y P G	D G L	G V S	S L R G
80	GATGAGCTGC	GGCTGGGACC TCGACGGCCT	CGGAGTTCTT	CGAGCCGGCA
1234567890	M S C	G W D L D G L	E F F	B P A
50 60 70 80 90 100	BsaBI CGATCGTCGC I V A	CACCGACCGC	CAGTTCGACG	GGGCCGGCAT
1234567890 1234567890 1234567890 1234567890		T D R	E F D A	A G I
60	BSABI GATCCGATCG CGATCGTCGC D P I A I V A	CGCCGTTCCC	CGACGCGGCC	GCCTTCGAGC
1234567890		P F P	D A A	A F E R
50	TGCCGACGAC	GAGGGCATCA	GGTTCCTGCA	GTCCTGGGAG
1234567890	A D D	E G I T	F L H	S W E
40	CCGCCACCGA	CGAGGGCGGC	CGCGAGGGCG	TCCTGACGAC
1234567890	G T D	E G G	R E G G	L T T
30 1234567890	GGCGCCGGCG	GGATGCTGTC M L S	GGCGTACGTC A Y V	GGACCCGCAG CAGCGGATGC TCC
10 20 30 40 1234567890 1234567890 1234567890	BSABI	AGCCCGGAG GACCTGTGGC G S P E D L W R	GACCCGGACG CGCTCGGCAG D P D A L G R	
10 1234567890	CGACTGTCGG T V G	CAGCCCGGAG S P E	GACCCGGACG D P D A	MSCI Ball • CGCTGGCCAT L A M

FIG. 31T

CGGTGTCTTC ATCGGCCTCT CCTACCAGGA CTACGCGGCC CGCGTCCGCG TGGCGTGGAG GGTTACCTGC TGACCGGCAG CACGCCGAGC 10000 G V F I G L S Y Q D Y A A R V P N A P R G V E G Y L L T G S T P S

6	1/	1	64
---	----	---	----

.0400	KpnI Acc65I ▼ ▼ A R R N G H P V L A V V R G T A V N Q D G A S N G L T A P N G P S Q	CGCGCCCAAC A P N	ACGGGCTGAC G L T	GGCGCCAGCA	CAACCAGGAC N Q D	KpnI Acc651 GGACCGCCGT TAV	Gregrecece C	GGTGCTCGCC GTGG	GCGCGGCGCA ACGGTCACCC	GCGCGGCGCA A R R N
0300	GCTCGCCCCC GACGCCCGA GCAAGGCCTT CTCGCGCCCCCCCCCC	TCGTGGAGCG V E R	CGGCGCCGC GGAGGGGGGTC GGCCTGCTGCTGGAGCG G A A E G V G L L L V E R	GGAGGGCGTC E G V	TCGGCGCCGC G A A	GCCGACGGGT TO	CTCGGCGAC S A D	GCAAGGCCTT K A F	ICTCGCCCG GACGGCCGCA GCAAGGCCT	GCTCGCCCG I A P
0200	Apali ** TGCGGGGCGCT GCGCAGCGCGCGCGCGCGCGTG GCGATGATGG CGACCCCGCA CATGTTCGTG GAGTTCAGCC GTCAGCGGGC 10200 RALRS G E C T M A L A G G V A M M A T P H M F V E F S R Q R A	CATGITCGIG GAGITCAGCC C	CATGTTCGTG M F V	GCGATGATGG CGACCCCGCA A M M A T P H	gcgatgatgg a m m a	CCTCGC CGGTGGCGTG	TGGCGCTCGC A L A	Apali GCCAGCGGC GAGTGCACGA R S G E C T M	GCGCAGCGGC R S G	TGCGGGCGCT R A L
0100	GTCGCGTCGG GCCGTATCGC GTACACCTTCGAAG GGCCGCGAC GACCGTCGAC ACCGCCTGCT CGTCGTCGCT GACCGCCCTG CACCTGGCGG 10100 V A S G R I A Y T F G L E G P A T T V D T A C S S S L T A L H L A V	GACCGCCTG T A L	CGTCGTCGCT S S L	GCCCGCGAC GACCGTCGAC ACCGCCTGCT CA P A T T V D T A C S	GACCGTCGAC T V D	GGCCCGCGAC P A T	GGTCTCGAAG G L E G	GTCGCGTCGG GCCGTATCGC GTACACCTTC GGTC V A S G R I A Y T F G L	GCCGTATCGC R I A	GTCGCGTCGG V A S G
	1234567890	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	100	06	80	70	09	20	40	30	20	10

FIG. 310

AGCAGCGGGT GATCCGGCAG GCGCTCGCCG ACGCCCGGC GACATCGACG CCGTCGAGAC GCACGCACG GGAACCTCGC TGGCCACC 10500 Q R V I R Q A L A D A R L A P G D I D A V E T H G T G T S L G D P

62/1	64
------	----

90 1234567890	CGGCCCGCGG AACGGCCGCT CGGGAAGT CCAACATCGG ACACACCCAG 10600 R P A E R P L A I G S V K S N I G H T Q	CGATGCGCCA CGGCACCTTCCACGCCGA CGAGCCGAGC	CGACTGGCCG GCCGCGCCG CGCCGCCGTC TCCTCCTTCG GCATCAGCGG 10800 D W P A G I G P R R A A V S S F G I S G	
50 60 70 80 90 100 1234567890 1234567890 1234567890 1234567890	TCCGTGAAGT CCAAC S V K S N	TCCACGCCGA CGAGC H A D E P	CGCCGCCGTC TCCTC	Bsu36I
0 0 1234567890	ACGGCCGCT CGCCATCGGC TCCGTGAAGT R P L A I G S V K S	GTLPKTLH AD	GCCGGCACCG GTCCGCGCCG A G T G P R R	
50 90 1234567890	3G AACGGCCGC E R P L	T G T L	CG GCCGGCACC	
30 90 12345678	TA CGGCAAGG! Y G K E	ATCATCAAG ATGGTCCTCG I I K M V L A	CA CCGAGCCG	
20 30 40 1234567890 1234567890	: CAGGGCCTCC AGGCCACGTA CGGCAAGGAG	O	CAGCGGCCTG GCCCTCGTCA CCGAGCCGAT S G L A L V T E P I	·
12345678	CAGGGCCT	GTGCGGCGGG	CAGCGGC	

FIG. 31V

3 GAACCTCCGA GGTCGCTGAG GGCTCTGAGG CCTCCGAGGC CCCGGCAGCC GTGAGGCGTC CCTCCCCGGG CACCTGCCCT 11000 3 T S E V A E G S E A S E A A P G S R R A S L P G H L P W

			63/164	
	23600	23700	23800	23900
100	CCGACGTGGC CCTGCCGGCCG GTGCCGAGGG CGCGCGGTTC GGCCTTCACC CGGCGCTGCT CGACGCCGCC GTGCAGGCGG CCGGTGCGGG 23600	CCGCGCCCTT CGCCGCGCGCTGC CGTTCGCCTG GAGCGGGATC TCCTGTACGC GGTCGGCGCC ACCGCCCTCC GCGTGCGGCT GGCCCCCGCC 23700	GGCCCGGACA CGGTGTCCGT GAGGCGCCGCC GGCAGCCGGT GTTCGCCGCG GACTCCCTCA CGGTGCTGCC CGTCGACCCC GCGCAGCTGG 23800	CGGCCTTCAG CGACCCGACT CTGGACTGCT GGAGTGGACC GCCTGGGACG GTGCCGGCGA GGCCCTGCCC GGCGCGGTCG TGCTGGGCGG 23900
1234567890	DVALPAVO A CPAG A GGAGG CGCGCGGTTC GCTTCACC CGGCGCTGCT CGACGCCGCC GTGCAGGCGG CCGGTGCGG 23600	R G V R R G H A A A V R L E R D L L Y A V G A T A L R V R L A P A	G P D T V S V S A A D S S G Q P V F A A D S L T V L P V D P A Q L A	A F S D P T L D A L H L L E W T A W D G A A Q A L P G A V V L G G
90 1234567890	CGACGCCGC GTGCAGGCGG (D A A V Q A A	GAGCGGGAIC ICCIGIACGC GGTCGGCGCC ACCGCCCTCC GCGTGCGGCT GGCCCCC E R D L L Y A V G A T A L R V R L A P	CGGTGCTGC CGTCGACCCC GCGCAGCTGG	ALPGA GCCCCCCTCCT
80	CGACGCCCC	ACCGCCCTCC	CGGTGCTGCC	GGCCTGCCC
1234567890	D A A	T A L R	V L P	A L P
70	CGGCGCTGCT	GGTCGGCGCC	GACTCCCTCA	FIGGACC GCCTGGGACG GTGCCGCGCA (
1234567890	A L L	V G A	D S L T	
60	GCCTTCACC	TCCTGTACGC	GTTCGCCGCG	GCCTGGGACG
1234567890		L Y · A	F A A	A W D G
50	CGCGCGGTTC	GAGCGGGATC	GGCAGCCGGT	GGAGTGGACC
	A R F	E R D L	Q P V	E W T
40	GTGCCGAGGG	CGTTCGCCTG	GACTCCTCCG	TGCACCTGCT
1234567890	A E G	V R L	D S S G	H L L
30	GAGGTCGCCG	ACGCGGCTGC	GAGCGCCGCC	CTGGACGCGC
1234567890	E V A G	A A A	S A A	L D A L
10 20 30 40 50 60 70 80 90 100 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	CARCGIGGC CCTGCCGGCC GTGCCGA(CGGGGCGTT CGGCGCGGCTGC R G V R R G H A A A	GGCCCGGACA CGGTGTCCGT GAGCGCCGCC GACTCCT(G P D T V S V S A A D S S	GCCTTCAG CGACCCGACT CTGGACGCGC TGCACCTG
1234567890	CCGACGTGGC D V A	CCGGGGCGTT R G V	GCCCGGACA GPDT	CGGCCTTCAG

FIG. 31W

CGACGCCGAC GGTCTCGCCG CGCCGCTGCG CGCCGAGGTCC TGTCCTTCCC GGACCTTACG GACCTGGTGG AGGCCGTCGA CCGGGGCGAG 24000 D A D G L A A A L R A G G T E V L S F P D L T D L V E A V D R G E

	11600	11700	11800	SphI 7 C 11900 L
1234567890	Sphi GOGOTOTICO GGOTOGICGA GAGOTGGGGGGGGGGGGGGGGGGGGGGGGGGCGCGGCGGCGGCG	CGCTCGCCGA CGCCCCCCC CTGGTCGCCG GCTCATGCAG GAGCTGCCCG CCGGTGGCGC GATGCTCGCC GTCCAGGCCG CGGAGGACGA 11700 L A D A A R L V A A R G R L M Q E L P A G G A M L A V Q A A E D E	GATCCGCCTG TGGCTGGAGA CGRACCCGGAGA CGTCTGGACG TCGCCGCCGT CAACGGCCCC GAGGCCGCCG TCCTGTCCGG CGACGCGGAC 11800 IRVWLETEE GORDACGG GTACCCCGGAC CGTCTGGACG TCGCCGCCG TCCTGTCCGG CGACGCGGAC 11800	Sphi GCGGCGCGG AGGCGGAGGC GTACTGGTC GGGCTCGGC GCGCTGCGG GTCAGCCACG CCTTCCACTC CGCGCACATG GACGGCATGC 11900 A A R E A B A Y W S G L G R R T R A L R V S H A F H S A H M D G M L
90	GCACGTCGCC	GTCCAGGCCG	TCCTGTCCGG	CGCGCACATG
1234567890	H V A	V Q A A	L S G	A H M
80	TCGCCGCCGC	GATGCTCGCC	GAGGCCGCCG	CCTTCCACTC
1234567890		M L A	B A A V	F H S
70 1234567890	Sphi ▼ ATGCGGCCGC GCACGTCACTCG GTCGGCGAGA TCGCCGCCGC GCACGTCGCC M R P A A L L G H S V G E I A A A H V A	GAGCTGCCCG CCGGTGGCGC GACCAGGCCG CGCA(ELPAGGCCG CGCACE)	CAACGGCCCC N G P	GTCAGCCACG V S H A
60	CGGTCACTCG	GAGCTGCCCG	TCGCCGCCGT	CGCGCTGCGG
1234567890	G H S	E L P A	A A V	A L R
50	CCGCACTGCT	GCGCGGCCG GCTCATGCAG	CGTCTGGACG	GCAGGACCCG
1234567890	A L L	R G R L M Q	R L D V	R T R
40 1234567890	Sphi A ATGCGGCCGG M R P A	CGCGCGGCCG R G R	GTACGCGGGA Y A G	G L G R
30	GAGCTGGGGC	CTGGTGGCG	CGGAGGAGCG	GTACTGGTCC
1234567890	S W G	L V A A	E E R	Y W S
10 20 30 40 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	GGCTCGTCGA	CGCCGCCGC CTGGTCG	TGGCTGGAGA	AGGCGGAGGC
	L V E	A A R L V A	W L E T	A B A
10 1234567890	GCGCTCTTCC GGCTCGTCGA	CGCTCGCCGA L A D	GATCCGCGTG I R V	GCGCCGCGG

pikPKS Sequence

FIG. 31X

TCGACGGGGTT CCGCGCCGTC CTGGAGAGGG TGGACCTCC CTGACCGTGG TCTCGAACGT CACCGGCCTG GCCGCCGGCC CGGACGACCT 12000

D G F R A V L E T V E F R R P S L T V V S N V T G L A A G P D D L

pikPKS Sequence

10	20 1234567890	10 20 30 40 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	40 1234567890	50 1234567890	60 1234567890	70 1234567890	80 1234567890	90 1234567890	1234567890	
TGCGACCCC	Scal GAGTACTGGG E Y W V	TCCGGCACGT R H V	CCGCGGCACC R G T	GICCGCTICC ICGACGGCGI V R F L D G V	TCGACGGCGT D G V	CCGTGTCCTG R V L	CGCGACCTCG R D L G	GCGTGCGGAC V R T	Scal ▼ GTGCGACCCC GAGTACTGGG TCCGGCACC GTCCGCTTCC TCGACGGCGT CCGTGTCCTG CGCGACCTCG GCGTGCGGAC CTGCCTGGAG 12100 C D P E Y W V R H V R G T V R F L D G V R V L R D L G V R T C L E	12100
TGGGCCCCG	ACGGGGTCCT G V L	CACCGCCATG T A M	GCGGCCGACG A A D G	GCCTCGCGGA L A D	CACCCCGCG	GATTCCGCTG D S A A	CCGGCTCCCC	CGTCGGCTCT V G S	CTGGGCCCCG ACGGGGTCCT CACCGCCATG GCGGCCGACG GCCTCCCCGCTG CCGGCTCCCC CGTCGGCTCT CCCGCCGGCT 122000 L G P D G V L T A M A A D G L A D T P A D S A A G S P V G S P A G S	12200
TCCCGCCGA P A D	CTCCGCCGCC S A A	GGCGCGCTCC	GGCCCGGCC	GCTGCTCGTG L L V	GCGCTGCTGC GC A L L R	GCCGCAAGCG R K R	GTCGGAGACC S B T	GAGACCGTCG E T V A	CTCCGCCGA CTCCGCCGCC GGCCCCGGCC GCTGCTCGTG GCGCTGCTGC GCGCAAGCG GTCGGAGACC GAGACCGTCG CGGACGCCT 12300 PADSAA GALRPRLIVALL STRKRSBTETVE DAL	12300
G R A	сасссссасс нав	GCACCGGACC T G P	ССАСТСССАС D W H	GCCTGGTTCG (A W F A	CCGGCTCCGG G S G	GGCGCACCGC A H R	GTGGACCTGC V D L P	CCACGTACTC T Y S	CGGCAGGGCG CACGCCCACG GCACTGGCAC GCCTGGTTCG CCGGCTCCGG GGCGCACCGC GTGGACCTGC CCACGTACTC CTTCCGGCGC 12400 G R A H A H G T G P D W H A W F A G S G A H R V D L P T Y S F R R	12400
GACCGCTACT D R Y W	GGCTGGACGC L D A	CCCGGCGGCC P A A	GACACCGCGG D T A V	TGGACACCGC D T A	CGGCCTCGGT G L G	CTCGGCACCG L G T A	CCGACCACCC D H P	GCTGCTCGGC L L G	GACCGCTACT GGCTGGACGC CCCGGCGGCG GACACCGCG CGGCCTCGGCACCG CCGACCACCC GCTGCTCGGC GCCGTGGTCA 12500 D R Y W L D A P A A D T A V D T A G L G L G T A D H P L L G A V V S	12500

FIG. 31Y

1 0 7

ם

LAFG

S N O

臼

G L

r D

> ⊠

POGA

M M

66/164

01	20	30	40	95	9	70	80	06	100	
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	1234567890	
GCCTTCCGGA L P D	CCGGGACGGC R D G	CTGCTGCTCA L L L T	CCGGCCGCCT G R L	CTCCCTGCGC	ACCCACCGT T H P W	GGCTCGCGGA L A D	CCACGCCGTC H A V	CTGGGGGAGCG L G S V	GCCTTCCGGA CCGGGACGGC CTGCTCA CCCTGCGC ACCCACCGT GGCTCGCGGA CCACGCGTC CTGGGGAGCG TCCTGCTCCC 12600 LPDRDBORDG LLLTGRLSLRTHPWLADHAV LGSVLLPP	12600
CGGCGCGCG G A A	ATGGTCGAAC M V E L	TCGCCGCGCA A A H	CGCTGCGGAG A A E	TCCGCCGGTC S A G L	TGCGTGACGT R D V	GCGGGAGCTG R E L	ACCCTCCTTG T L L E	AACCGCTGGT P L V	CGGCGCCGCG ATGGTCGAAC TCGCCGGAG TCCGCCGGTC TGCGTGACGT GCGGGAGCTG ACCCTCCTTG AACCGCTGGT ACTGCCCGAG 12700 G A A M V E L A A H A A E S A G L R D V R E L T L L E P L V L P E	12700
CACGGTGGCG H G G V	CACGGTGGCG TCGAGCTGCG H G G V E L R	CGTGACGGTC V T V	GGGGCGCCGG	CCGGAGAGCC G E P	CGGTGGCGAG G G E	TCGGCCGGGG S A G D	ACGGCGCACG G A R	GCCCGICICC P V S	CACGGREGCG TCGAGCTGCG CGTGACGGCCCGG CCGGAGAGCC CGGTGGCGAAG TCGGCCGGGG ACGGCGCACG GCCCGTCTCC CTCCACTCGC 12800 H G G V E L R V I V G A P A G E P G G E S A G D G A R P V S L H S R	12800
GGCTCGCCGA	CGCGCCCGCC	Kpni Acc651 W W GGTACCGCCT G I A W	GGTCTGCCA S C H	CGCGACCGGT A T G	MSCI Ball CTGCTGGCCA (I I CCCACCGGCC D R P	CGAGCTTCCC E L P	GTCGCGCCCG V A P D	KpnI Acc651 Ball ▼ ▼ ■ GGCTCGCCGC GGTACCGCCT GGTCCTGCCA CGCGACCGGCC CGAGCTTCCC GTCGCGCCCG ACCGTGCGGC 12900 L A D A P A G T A W S C H A T G L L A T D R P E L P V A P D R A A	12900
CATGTGGCCG	בכפכאפפפכפ	CCGAGGAGGT	GCCGCTCGAC	GGTCTCTACG	AGCGGCTCGA	CGGGAACGGC	CTCGCCTTCG	GTCCGCTGTT	CATGIGGCCG CCGCAGGGGG CCGAGGAGGT GCCGCTCGAC GGTCTTACG AGCGGCTCGA CGGGAACGGC CTCGCCTTCG GTCCGCTGTT CCAGGGGCTG 13000	13000

pikPKS Sequence

FIG. 31Z

67/	164
-----	-----

10	20	30	40	50	09	70	80	06	100	
1234567890	1234567890 1234567890 1234567890 1234	1234567890	1234567890	1234567890	1234567890	567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	
AACGCGGTGT G N A V W	GGCGGTACGA R Y E	GGGTGAGGTC TTCG	TTCGCCGACA F A D I	TCGCGCTCCC A L P	CGCCACG A T T	BSMI CGACA TCGCGCTCCC CGCCACCACG AATGCGACCG CGCCCGCGAC CGCCGGGAGTG D I A L P A T T N A T A P A T A N G G G S A	CGCCCGCGAC P A T	CGCGAACGGC A N G	Bemi AACGCGGTGT GGCGGTACGA GGGTGAGGTC TTCGCCGACA TCGCGGTCCC CGCCACCACCACG AATGCGACCG CGCCCGCGACCGC GGCGGGAGTG 13100 N A V W R Y E G E V F A D I A L P A T T N A T A P A T A N G G G S A	13100
CGGCGGCGGC A A A	A A A P Y G I H P A	ATCCACCCG I H P A	CCCTGCTCGA L L D	CGCTTCGCTG A S L	CACGCCATCG H A I A	CACGCCATCG CGGTCGGCGG TCTCGTCGAC GAGCCCGAGC	TCTCGTCGAC L V D	GAGCCCGAGC E P E L	CGGCGGCGGC CCCTACGGC CCCTGCTCGA CGCTTCGCTG CACGCCATCG CGGTCGGCCG TCTCGTCGAC GAGCCCGAGC TCGTCCGCGT 13200 A A A P Y G I H P A L L D A S L H A I A V G G L V D E P E L V R V	13200
CCCCTTCCAC P F H	TGGAGCGGTG WSGV	Apali TGGAGCGGTG TCACCGTGCA C W S G V T V H	I CGCGGCCGGT A A G	GCCGCGGCGG C	CCCGGGTCCG R V R	RVR LAS	S GCGGGGACGG ACGCCGTCTC	ACGCCGTCTC A V S	Apali * CCCCTTCCAC TGGAGCGGTG TCACCGTGCTGCTGCTCC GCGGGGACGG ACGCCGTCTC GCTGTCCCTG 13300 PFHWSGVTT VHAAGABACGTG ACGCGGCGGCGGGGGTCCG TCTCGCCTCC GCGGGGACGG ACGCCGTCTC GCTGTCCCTG 13300	13300
ACGGACGGCG AC T D G E	AGGACGCCC GRP	GCTGGTCTCC L V S	GTGGAACGGC T V E R L	TCACGCTGCG T L R	CCCGGTCACC P V T	GCCGATCAGG C	CGGCGGCGAG A A S	CCGCGTCGGC R V G	ACGGACGGCG AGGGACGCCC GCTGGTCCC GTGGAACGGC TCACGCTGCG CCCGGTCACC GCCGATCAGG CGGCGGCGAG CCGCGTCGGC GGGCTGATGC 13400 T D G E G R P L V S V E R L T L R P V T A D Q A A A S R V G G L M H	13400

FIG. 31AA

ACCGGGTGGC CTGGCGTCCG TACGCCCTCG CCTCGTCCGG CGCACGCCC CTCCTACGG GCCGACCGCC GTCCTCGGCA AGGACGAGCT 13500 R V A W R P Y A L A S S G B Q D P H A T S Y G P T A V L G K D E L

1234567890	10 20 30 40 50 60 70 80 90 1034567890 12346780 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 12345678	30 1234567890	40 1234567890	50 1234567890	60 1234567890	70 1234567890	80 1234567890	90 1234567890	100 1234567890	
		いていている。	し むで & でしむし	しては出いませてい	したび止して をひこと	Constant	CAGGACGTGG	ວຍດອຍດວຍຄວ	0.0551 335035 3555555555555555555555555555	3600
K V A	K V A A A L E	S A G	М М М М М М М М М М	4 T D	U A	S I	EVGLYPDLAALSQDVAAGA	е В	4 d	
CGTACCGTCC 1 R T V L	Trgcgccgcr A P L	GCCCGCGGGT P A G	CCCGCCGACG P A D G	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	GGGTGTACGG G V R	GGCACGGTGG G T V A	GGGTGTACGG GGCACGGTGG CCCGGAGCTGCTC CAGGCCTGGC G V R G T V A R T L E L L Q A W L	GGAGCTGCTC E L L	CGTACCGTCC TTGCGCCGCT GCCCGCGGGTC GCGCCCGGA GGGTGTACGG GGCACGGTGG CCCGGACGCT GGAGCTGCTC CAGGCCTGGC 13700 R T V A R T L E L L Q A W L	3700
TGGCCGACGA (A D E Not	A GCACCTCGCG H L A Noti	GCACCGCC G T R L	recreers r r v	CACCCGCGGT T R G	GCGGTGCGGG A V R D	ACCCCGAGGG P B G	CTCCTGGT CACCCGCGGT GCGGTGCGGG GTCCGGCGCC GACGATGGCG GCGAGGACCT L L V T R G A V R D P B G S G A D D G G B D L	Gacgatggcg D D G G	TGGCCGACGA GCACCTCGCG GGCACCCGCC TGCTCCTGGT CACCCCGCGGG ACCCCGAGGG GTCCGGCGCC GACGATGGCG GCGAGGACCT 13800 A D E H L A G T R L L L V T R G A V R D P B G S G A D D G G E D L Noti	3800
GTCGCACGCG S H A	GCCCCTGGG A A W G	GTCTCGTACG L V R	GACCGCGCAG T A Q	ACCGAGAACC T E N P	ACCGAGAACC CCGGCCGCTT TENPGRE	CGGCCTTCTC G L L	CGGCCTTCTC GACCTGGCCG ACGACGCCTC G L L D L A D D A S BStXI	ACGACGCCTC D A S	GICGCACCCC GCCCCTGGG GICTCGTACG GACCGCGAGAACC CCGGCCGTT CGGCCTTCTC GACCTGGCCG ACGACGCCTC GTCGTACCGG 13900 SHAAAWG LVR TAQ TENPGRFFGLLDLADDASS SYR BBtXI	3900
ACCCTGCCGT T L P S	ACCCTGCCGT CGGTGCTCTC CGACGCGGG	CGACGCGGGC D A G	CTGCGCGACG	AACCGCAGCT P Q L	CGCCCTGCAC A L H	GACGGCACCA D G T I	TCAGGCTGGC R L A	CCGCCTGGCC R L A	ACCCTGCCGT CGGTGCTCTC CGACGCGGGC CTGCGCGACG AACCGCAGCT CGCCCTGCAC GACGGCACCA TCAGGCTGGC CCGCCTGGCC TCCGTCCGGC 14000 T L P S V L S D A G L R D E P Q L A L H D G T I R L A R L A S V R P	4000

pikPKS Sequence

FIG. 31BB

69/	16	4
-----	----	---

CGGTCCTG CTGACCGGCG GCACCGGCG	Pmli Y CCGAGACCGG CACCGCAGGG CACGGTCCTG CTGACCGGCG GCACCGGCGG CCTGGGCGGA CTGGTCGCCC GGCACGTGGT 14100 B T G T A A P A L A P E G T V L L T G G T G G L G G L V A R H V V
GGGCACGG ACGCCCGGG CGCCGACCG T D A P G A D E	Apali GGGCGAGTGG GGCGTACGAC GCTGCTGCT GGTGAGCCGG CGGGGCACGGG CGCCGACGAG CTCGTGCACG AGCTGGAGGC CCTGGGAGCC 14200 G E W G V R R L L L V S R R G T D A P G A D E L V H E L E A L G A
CTCACCGC CGTACTCGAC GCCATCCC L T A V L D A I P	GACGICTCGG TGGCCGCGG CGACGCGCGAAG CCCTCACCGC CGTACTCGAC GCCATCCCCG CCGAACACCC GCTCACGCG GTCGTCCACA 14300 D V S V A A C D V A D R E A L T A V L D A I P A E H P L T A V H T
CGGAGGAC GTGGAACACG TACTGCG	CGGCAGGCGT CCTCTCCGAC GGCCCTCC CGTCCATGAC GACGGAGGAC GTGGAACACG TACTGCGGCC CAAGGTCGAC GCCGCGTTCC TCCTCGACGA 14400 A G V L S D G T L P S M T T E D V E H V L R P K V D A A F L L D E
GITCTCCT CCGCCGCCGC CGTCTTCGGT GGCGCGGGC AGGGCGCCTA	ACTCACCTCG ACGCCCGCAT ACGACCTGGC AGGGTTCGTC CCGCCGCCGC CGTCTTCGGT GGCGCGGGGC AGGGCGCTA CGCCGCCGCC 14500 L T S T P A Y D L A A F V M F S S A A N V F G G A G Q G A Y A A A

pikPKS Sequence

70/	1	64
-----	---	----

10	10 20 30 40 50 60 70 80 90 100 100 100 1034567890 123456780 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 123456	30	40	50	60	70	80 1234567890	90 1234567890	1234567890	
060/06/571	000000000000000000000000000000000000000									
AACGCCACCC N A T L	AACGCCACCC TCGACGCCCT N A T L D A L	CGCCTGGCGC A W R	CGCCGGGCAG R R A A	CCGGACTCCC G L P	G L P A L S	CTCGGCTGGG GCCTCTGGGC	GCCTCTGGGC L W A	CGAGACCAGC E T S	AACGCCACC TCGACGCCT CGCCTGGCGC CGCCGGGCAG CCGCCTCTCGCTGGG GCCTCTGGGC CGAGACCAGC GGCATGACCG 14600 N A T L D A L A W R R R A A G L P A L S L G W G L W A B T S G M T G	14600
GCGAGCTCGG E L G	GAGCTCGG CCAGGCGGAC CTGCGCCGGA E L G Q A D L R R M	CTGCGCCGGA L R R M	TGAGCCGCGC S R A	GGCATCGGC G I G	GGGATCAGCG A	ACGCCGAGGG A B G	CATCGCGCTC I A L	CTCGACGCCG L D A A	GCGAGCTCGG CCAGGCGGAC CTGCGCCGGG TGAGCCGCGCGCGCGGCGCG	14700
CGACCGCCAC D R H	CCGGTCCTGC TGCCCCTG P V L L P L	TGCCCTGCG P L R	GCTCGACGCC L D A	GCCGGGCTGC A G L R	GGGACGCGC D A A	CGGGAACGAC G N D	CGGGAACGAC CCGGCCGGAA G N D P A G I	TCCCGGCGCT P A L	CGACCGCCAC CCGGTCCTGC TGCCCCTGCG GCCGGGTGC GGGACGGC CGGGAACGAC CCGGCCGGAA TCCCGGCGCT CTTCCGGGAC 14800 DRHPVLLPVLLPPA BAGLRDA AG COMBANA GA BAGIPA GIPAL FRD	14800
GTCGTCGGCG (CCAGGACCGT	CCGGGCCCGG R A R	CCGTCCGCGG P S A A	CCTCCGCCTC S A S	GACGACAGCC T T A	GGGACGGCCG GC G T A G	GCACGCCGGG TPG	GACGGCGGAC T A D	GTCGTCGGCG CCAGGACCGT CCGGCCCGG CCTCCGCCTC GACGACAGCC GGGACGGCCG GACGCCGGG GACGGCGGAC GGCGCGGCGG 14900 V V G A R T V R A R P S A S A S T T A G T A G T P G T A D G A A E	14900
AAACGGCGC (GGTCACGCTC	GCCGACCGGG A D R A	CCGCCACCGT A T V	GGACGGGCC D G P	GCACGCAGC A R Q R	S GCACGCCAGC GCCTGCTGCT A R Q R L L L L	Xhoi PaeR7I r cgagrīcgrc E F V	GTCGGCGAGG V G E V	XhoI PaeR7I ▼ ▼ T A A V T L A D R A A T V D G P A R Q R L L L E F V V G E V A E V	15000

7	1/	1	64
---	----	---	----

		•	.,		
	15100	15200	15300	15400	15500
1234567890	Scal ACTCGGCCAC GCCCGCGGATCGA CGCCGAACGG GGCTTCCTCG ACCTCCCTCG ACCGCCGTCG AACTCCGCAA CCGGCTCAAC 15100 L G H A R G H R I D A B R G F L D L G F D S L T A V B L R N R L N L G H A R G H R I D A B R G F L D L G F D S L T A V B L R N R L N	TCCGCCGGTG GCCTCGCCCT CCCGGCGTCTTCG ACCACCCAAG CCCGGCGGCA CTCGCCTCCC ACCTGGACGC CGAGCTGCCG CGCGGCGCT 15200 S A G G L A L P A T L V F D H P S P A A L A S H L D A E L P R G A S	CGGACCAGGA CGGAGCGGG AACCGGAACGAGAA CGGGACGACG GCGTCCCGGA GCACCGCCGA GACGGACGCG CTGCTGGCAC AACTGACCCG 15300 D Q D G A G N R N G N E N G T T A S R S T A E T D A L L A Q L T R	CCTGGAAGGC GCCTTGGTGC TGACGGGCCTCTCGGGACGG AAGAAGTCCT GGAGCACCTG CGGTCCCTGC GCTCGATGGT CACGGGCGAG 15400 L E G A L V L T G L S D A P G S E E V L E H L R S L R S M V T G E	ACCGGGACCG GGACCGCCTC CGGAGCCCCG GGTCCGGCGC CGAGGACCGG CCCTGGGGGACGG CCGGGGACGG AGCCGGGGGC GGGAGTGAGG 15500 T G T G T A S G A P D G A G S G A B D R P W A A G D G A G G G S B D
90 1234567890	AACTCCGCAA L R N	CGAGCTGCCG E L P	SACGGACGCG CTGCTGGCAC T D A L L A Q	GCTCGATGGT S M V	AGCCGGGGGC A G G
40 50 60 70 80 90 100 567890 1234567890 1234567890 1234567890 1234567890 1234567890	GGCTTCCTCG ACCTCGGTT CGACTCCCTG ACCTCCGCAA G F L D L G F D S L T A V E L R N	TTCG ACCACCAAG CCCGGCGGCA CTCGCCTCCC ACCTGGACGC CGAGCTGCCG	GACGGACGCG T D A	CGGTCCCTGC R S L R	CCGGGGACGG
70 1234567890	CGACTCCCTG D S L	CTCGCCTCCC L A S H	GCACCGCCGA T A E	GGAGCACCTG E H L	CCCTGGGCGG
60 1234567890	ACCTCGGCTT L G F	CCCGGCGGCA P A A	GCGTCCCGGA GCACCGCCGA A S R S T A E	AAGAAGTCCT E V L	CGAGGACCGG E D R
50 1234567890	GGCTTCCTCG	ACCACCCAAG H P S	CGGGACGACG G T T	CCCGGGAGCG	GGTCCGGCGC S G A
40 1234567890	CGCCGAACGG A E R	CTGGTCTTCG	ggaacgagaa n e n	CTCGGACGCC S D A	GACGGCGCCG D G A G
30	ACCGGATCGA R I D	CCCGGCGACC P A T	AACCGGAACG N R N G	TGACGGGCCT T G L	CGGAGCCCCG
10 20 30 1234567890 1234567890 1234	GCCCGCGGTC ACCGGATCG? A R G H R I D	GCCTCGCCT L A L	CGGAGCCGGG G A G	GCCTTGGTGC A L V L	GGACCGCGTC T A S
1234567890	Scal ▼ ACTCGGCCAC L G H	TCCGCCGGTG GCCTCGCCCT S A G G L A L	CGGACCAGGA D Q D	CCTGGAAGGC L E G	ACCGGGACCG GG T G T G

10	20	30	. 40	50	09	70	80	06	100	
1234567890	1234567890 1234567890 1234567890 1234	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	
ACGGCGCGGG	BBPHI ** ACGGCGCGGG AGTGCCGGAC TICATGAACG CCTCGGCCGA GGAACTCTIC GGCCTCCTCG ACCAGGACCC CAGCACGGAC TGATCCCTGC CGCACGGTCG 15600	BSPHI TTCATGAACG	CCTCGGCCGA	GGAACTCTTC	GGCCTCCTCG	ACCAGGACCC	CAGCACGGAC	TGATCCCTGC	CGCACGGTCG	15600
ស ស ស	О Ф	V P O G V	∝ √	E L F Xmm1	B L F G L L D Q D P Xmm1	a 0	O T O	ָרָ בְּיָבְיָבָיִרְ בְּיִבְּיִבְיִיבְיִיבְיִיבְיִיבְיִיבְיִי	ָרֶ בְּבָּיִבְיִבְּיִיבְיִיבְיִיבְיִיבְיִיבְיִיבְ	, C C C U
CCTCCCGCCC	CCTCCGCCC CGGACCCCGT CCGGGCACC TCGACTCGAA TCACTTCAIG CGCCCTCCGG GCGCCTCCAG GAACTCAAGG GSALAGCGG TCCACGGGA 13/00 V S T V N	CCCGGGCACC	TCGACTCGAA	TCACITCALG	990.I009090	ececci ccae	GAACTCHAGG	V V	S T V N	00/61
acgaagaa e e k	ACGAAGAAA GTACCTCGAC TACCTGCGTC GTGCCACGCCCCCCCCCC	TACCTGCGTC Y L R R	GTGCCACGGC A T A	GGACCTCCAC D L H	GAGGCCCGTG E A R G	GAGGCCGTG GCCGCCTCCG E A R G R L R	CGAGCTGGAG E L E	GCGAAGGCGG GCGAGCCC	GCGAGCCGGT E P V	15800
GGCGATCGTC A I V	GGCGATCGTC GGCATGGCCT GCCGCCGCGCGCGTC GCCTCGCCG AGGACCTGTG GCGGCTGGTG GCCGGCGGCG AGGACGCGAT CTCGGAGTTC 15900 A I V G M A C R L P G G V A S P E D L W R L V A G G E D A I S E F	GCCGCCTGCC R L P	CGGCGGCGTC G G V	GCCTCGCCCG A S P E	GCCTCGCCCG AGGACCTGTG A S P E D L W	GCGGCTGGTG R L V	GCCGGCGGCG	GCGGCTGGTG GCCGGCGCGCGAT R L V A G G E D A I	CTCGGAGTTC S E F	15900

pikPKS Sequence

FIG. 31FF

CCCCAGGACC GCGCTGGGA CGTGGAGGGC CTGTACGACC CGAACCCGGA GGCCACGGC AAGAGTTACG CCGCGAGGC CGGATTCCTG TACGAGGCGG 16000 P Q D R G W D V B G L Y D P N P E A T G K S Y A R E A G F L Y E A G

_
w
O
~
=
Φ
-3
=
Ų
•
Ŵ
•
40
S
\sim
_
₽.
$\overline{}$
=
\overline{a}
-

			73/164		ľ
	16100	16200	16300	16400	16500
1234567890	AGGCGTTCGA A F E	Bamhi ▼ GCACGCCGG AICCCGGCGG CCACCGCGCG GCGCGCCTCT TCACCGGCGT GATGTACCAC GACTACGCCA CCCGTCTCAC CGATGTCCCG 16200 HAGIPAAT RIPAT TARGIS VGVFIGV	GAGGGCATCG AGGGCTACCT GGGCACCCCCCCCCCCCC	ACACCGCCTG CTCGTCCTCGCCC TGCACCTCGC CGTGCAGGCC CTCCGCAAGG GCGAGGTCGA CATGGCGCTC GCCGGCGGCG TGACGGTCAT 16400 T A C \$ S S L V A L H L A V Q A L R K G E V D M A L A G G V T V M	XmnI ▼ GTCGACGCC AGCACCTTCG TCGACTAGCGC GGGCTGGCGC CGGACGGCCG GTCGAAGTCC TTCTCGTCGA CGGCCGACGG CACCAGCTGG 16500 S T P S T F V E F S R Q R G L A P D G R S K S F S S T A D G T S W
90 1234567890	GCCTCCTGGG A S W E	GTCGGCGTCT TCACCGGCGT GATGTACCAC GACTACGCCA CCCGTCTCAC CGATGTC V G V F T G·V M Y H D Y A T R L T D V	CA GTGTCGCCTC GCGTACACGC TTGGCCTGGA GGGGCCGGCC GTCACGGTCG S V A S G R V A Y T L G L E G P A V T V D	CGC CGTGCAGGCC CTGCGCAAGG GCGAGGTCGA CATGGCGCTC GCCGGCGGCG TGACGGTCA	CGGCCGACGG A D G
80 1234567890	CCTCCTGGAG L L E	GACTACGCCA D Y A T	TTGGCCTGGA G L E	CATGGCGCTC M A L	nI TTCTCGTCGA F S S T
70 1234567890	AGCAGCGICT Q R L	GATGTACCAC M Y H	GCGTACACGC A Y T L	GCGAGGTCGA E V D	XmnI GTCGAAGTCC T' S K S F
60 1234567890	ATGGACCCGC M D P Q	TCACCGGCGT T G. V	GGGCGCGTC G R V	CTGCGCAAGG L R K G	CGGACGGCCG D G R
50 1234567890	GGCCCTCGCC A L A	GTCGGCGICT V G V F	GTGTCGCCTC V A S	CGTGCAGGCC V Q A	GGGCTGGCGC G L A P
40	CGCCGCGCGA P R E	CGGCACCTCG G T S	AACTCCGGCA N S G S	TGCACCTCGC H L A	CCGTCAGCGC R Q R
30	TTCGGGATCT F G I S	CCACCGCGCG	GGCACCGGC G T G	CTGGTCGCCC TGCACCT	TCGAGTTCAG E F S
20	CGCCGACTTC	BamHI FICACGCCGGG ATCCCGGCGG CCACCGCGCG H A G I P A A T A R	G I E G Y L	CTCGTCCTCG \$ S S	AGCACCTTCG S T F V
10 20 30 40 50 60 70 80 90 100 100 100 100 1234567890 123456780 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 12	GCGAGTICGA CGCCGACTIC TICGGGATCT CGCCGCGA GGCCTCGCC ATGGACCCGC AGCAGCGICT CCTCCTGGAG GCCTCCTGGG AGGCGTTCGA 16100	BamHI GCACGCGGG AI H A G I	GAGGGCATCG E G I E	ACACCGCCTG T A C	GTCGACGCCC S T P

œ Qi. മ്പ

	7	4	/	16	4
--	---	---	---	----	---

1	_	,	7/ 107	٥	0
	16600	1670(1680(1690	1700
1234567890	BamHI TCCGAGGGCG TCGGCGTCGTCGAG CGCCTGTCCG ACGCGCGTCG CAAGGGCCAT CGGATCCTCG CCGTGGTCCG GGGCACCGCC GTCAACCAGG 16600 S E G V G V L L V E R L S D A R R K G H R I L A V V R G T A V N Q D	ACGGCGCCAG CAGCGGCCTC CGAGCGCGTC GCAGCGC GTCATCCGAC GTGCCCTGGC GGACGCCCGG CTCACGACCT CCGACGTGGA 16700 G A S S G L T A P N G P S Q Q R V I R R A L A D A R L T S D V D	CGICGICGACGA CGGGIACGCG ACTCGGCGAC CCGAICGAGG CGCAGGCCGI CAICGCCACG IACGGGCAGG GCCGIGACGG CGAACAGCCG 16800 V V B A H G I G I R L G D P I E A Q A V I A I Y G Q G R D G E Q P	Beli ** CTGCGCCTCG GGTCGTTGAA GTCCAACATC GGACACCCCCC CGGTGTCTCC GGCGTGATCA AGATGGTCCA GGCGATGCGC CACGGCGTCC 16900 L R L G S L K S N I G H T Q A A A G V S G V I K M V Q A M R H G V L	Pmli ▼ TGCCGAAGAC GCTCCACGTG GAGAAGCCGA CGGACCAGGT GGACTGGTCC GCGGCGCGG TCGAGCTGCT CACCGAGGCC ATGGACTGGC CGGACAAGGG 17000
90	GGGCACCGCC	CTCACGACCT	GCCGTGACGG	GGCGATGCGC	ATGGACTGGC
1234567890	G T A	L T T S	R D G	A M R	
80 1234567890	CCGTGGTCCG V V R	GGACGCCGG D A R	CCGATCGAGG CGCAGGTCATCGCCACG TACGGGCAGG GCCGTGACGG	AGATGGTCCA M V Q	CCAGGT GGACTGGTCC GCGGGCGGG TCGAGCTGCT CACCGAGGCC ATGGACTGGC CGGACAAGG
70 1234567890	BamHI CGGATCCTCG R I L A	GTGCCCTGGC A L A	CATCGCCACG I A T	Beli GGCGTGATCA G V I K	TCGAGCTGCT
60	CAAGGGCCAT	GTCATCCGAC	CGCAGGCCGT	CGGTGTCTCC	5505055555
1234567890	K G H	V I R R	Q A V	G V S	550555555555555555555555555
50 1234567890	ACGCGCGTCG A R R	GCAGCAGCGC Q Q R	CCGATCGAGG P I E A	TO A A A	GGACTGGTCC
40	CGCCTGTCCG	ACGGGCCGTC	ACTCGGCGAC	C GGACACACCC	CGGACCAGGT
1234567890	R L S D	G P S	L G D	G H T Q	
30	CCTCGTCGAG	ACGGCTCCGA	CGGGTACGCG	GTCCAACATC	GAGAAGCCGA
1234567890	L V E	T A P N	G T R	S N I	
20	TCGGCGTCCT	CAGCGGCCTC	GTCGTCGAG GCCCACGGCA	CTGCGCCTCG GGTCGTTGAA GTCCAACAT	PM11 ▼ FCGAAGAC GCTCCACGTG GAGAAGCCGA (
1234567890	G V L	S G L	V V B A H G T	LRLGSLKSNI	
10 20 30 40 50 60 70 80 90 100	TCCGAGGGCG T	ACGGCGCCAG	CGTCGTCGAG	CTGCGCCTCG	TGCCGAAGAC
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	S E G V	G A S	V V B	L R L G	

FIG. 31HH

75/164

		•		
17100	17200	17300	17400	17500
GGAGACCCCT	GCCCAGATCG	TCGAGCACCG	CCGGGTGGCG	TGCGAGAGCG
E T P	A Q I G	E H R	R V A	C E S A
CGGCGGCCGA (CGCGCTGGAC A L D	ABCI ▼ CGCGCCGAGT R A E F	CGGACGTGGG D V G	CATGGCCGAG M A E
GAAGAGCCC E E A P	AGACTCCGGC T P A	GGCCGGCGGG	GCACGCCT G T P S	XmnI ▼ WATGG GCGCCGAACT CCTCGACGTG TCGAAGGAGT TCGCGGCGGC CATGGCCGAG TGCGAGAGCG W G A E L L D V S K E F A A A M A E C E S A
CGTCGTGCTC V V L	GTGTCGGCGA V S A K	CTCGCGTACT R V L	ACTGATACGC L I R	XmnI TCGAAGGAGT S K E F
CGAACGCGCA	GCCGTGGCTG	GCCCCGTCG	CTCCGGAAGG	CCTCGACGTG
N A H	P W L		P E G	L D V
GTCAGCGGGA	CCGGCCTGGT	CGCCGATCCG	GCGCTGACCG	GCGCCGAACT
V S G T	G L V	A D P	A L T A	A E L
CTCCTTCGGC	TCGGTCGGCG	GTACGGACGC	TTTCGCGCAG	GCCGGGATGG
S F G	S V G A	T D A	F A Q	A G M G
CCGCGGTCTC	CGTCGAGCCG	TCGCAGGGCC	Gacaggacga	CACGCAGTGG
A V S	V E P	S Q G R	Q D D	T Q W
CTGCGCAGGG	CGACCCCGGC	CGCGTTCGCC	CTCGGCACCG	XmnI * TTCGTGTTCC CCGGTCAGGG CACGCAGATGG GCGCCGAACT CCTCGACGTG TCGAAGGAGT TCGCGGCGGC CATGGCCGAG TGCGAGAGCG 17500 P V P P G Q G T Q W A G M G A E L L D V S K E F A A M A E C E S A
L R R A	T P A	A F A	L G T G	
CGACGGCGGA	GCCTCCGAGG	GACGCCTCGC	GGCCGTCGTG	TTCGTGTTCC
D G G	A S B A	R L A	A V V	F V F P
	TCGGC GTCAGCGGA CGAACGCGA CGTCGTGCTC GAAGAGGCCC CGGCGGCCGA (CGACGCCGA CTGCGCGTCTC CTCCTTCGGC GTCAGCGGGA CGAACGCGCA CGTCGTGCTC GAAGAGGCCC CGGCGGCCGA GGAGACCCCT 17100 D G G L R R A A V S S F G V S G T N A H V V L E E A P A A E E T P GCCTCCGAGG CGACCCCGGC CGGCCTGGT GCCGTGGCTG GTGTCGGCGA AGACTCCGGC CGCCTGGAC GCCCAGATCG 17200 A S E A T P A V E P S V G A G L V P W L V S A K T P A A L D A Q I G	CANCESCUER CONTROLLE CTCCTTCGGC GTCAGCGGGA CGACGCGCA CGTCGTGCTC GAAGAGGCCC CGGCGCCGA GGAGCCCT 17100 D G G L R R A A V S S F G V S G T N A H V V L E E A P A B E T P GCTCCGAGG CGACCCCGGC CGCGCTGGCG CCGCCTGGT GCCGTGGCTG GTCTCGGCGA AGACTCCGGC CGCGTGGAC GCCCAGATCG 17200 A S E A T P A V E P S V G A G L V P W L V S A K T P A A L D A Q I G A S E A T P A V E P S V G A G L V P W L V S A K T P A A L D A Q I G GACCCTCGC CGCGTTCGCC TCGCAGGGCC GTACGGACCG GCCCGATCCG GCCCGAGTCG GCCCGAGT TCGAGCACCG 17300 R L A A F A S Q G R T D A A D P G A V A R V L A G G R A E F E H R	CANCECCRAGA CTGCGCAGGG CCCCCGTCGGC GTCAGCGGGA CGAACGCGCA CGTCGTGCTC GAAGAGGCCC CGGCGGCCGA GGAGACCCCT 17100 G G L R R A A V S S F G V S G T N A H V V L B E A P A B E T P A S E A T P A V E P S V G A G L V P W L V S A K T P A A L D A Q I G GACCCTCGAGG CGCCTCGAGGC CGCCGATCCG GCCGCGTCG CTCGCGTACTCCGGC CGCGCTGGAC GCCCTGATCG 17200 A S E A T P A V E P S V G A G L V P W L V S A K T P A A L D A Q I G GACCCTCGC CGCTTCGC CGCTACGGACG CGCCGATCCG GCGCGGTCG CTCGCGTACT GGCCGGAGT TCGACCACCG 17300 R L A A F A S Q G R T D A A D P G A V A R V L A G G R A E F E B R GGCCGTCGTG CTCGCACCG GACAGGACGA GCGCTGACCG CTCCGGAAGG ACTGATACGC GGCACGCCCT CGGACGTGGC CCGCGTGGC CTCGGGTGGCG 17400 A V L G T G Q D D F A Q A L T A P E G L I R G T P S D V G R V A

76/	1	64
-----	---	----

			•	•	
	17600	17700	17800	17900	18000
1234567890	ABCI CGCTCTCCCCG CTATGTCGAC TGGTCGCTGG AGGCCGTCGT CCGGCAGGCG CCGGGCGC CCACGCTGGA GCGGGTCGAC GTCGTCCAGC CCGTGACCTT 17600 L S R Y V D W S L E A V V R Q A P G A P T L E R V D V V Q P V T F	CGCTGTCAIG GITTCGCTGG CGAAGGICTG GCAGCACCAC GGCGTGACGCCGT CGTCGGCCAC TCGCAGGGCG AGATCGCCGC CGCGTACGTC 17700 A V M V S L A K V W Q H H G V T P Q A V V G H S Q G E I A A Y V	GCCGGTGCCC TCACCCTCGA CGACGCCGCC CGCGTCGTCA GCCGCCCACC TCGCCGGCAA GGGCGGCATG ATCTCCCTCG 17800 A G A L T L D D A A R V V T L R S K S I A A H L A G K G G M I S L A	CCCTCAGCGA GGAAGCCACC CGGCAGCGA TCGAGAACCT CCACGGACCG CCGTCAACGG CCCCACCGCC ACCGTGGTTT CGGGCGACCC 17900 L S E E A T R Q R I E N L H G L S I A A V N G P T A T V V S G D P	CACCCAGATC CAAGAGGTGG CTCAGGCGGTG GGGGTCCGCG CACGGATCAT CCCCGTCGAC TACGCCTCCC ACAGCGCCCCA CGTCGAGACC 18000 T Q I Q E L A Q A C E A D G V R A R I I P V D Y A S H S A H V E T
90	GTCGTCCAGC	AGATCGCCGC	GGGCGGCATG	ACCGTGGTTT	ACAGCGCCCA
1234567890	V V Q P	I A A	G G M	T V V S	S A H
80	GCGGGTCGAC	TCGCAGGGCG	TCGCCGGCAA	CCCCACCGCC	TACGCCTCCC
1234567890	R V D	S Q G E	A G K	P T A	Y A S H
70	CCACGCTGGA	CGTCGGCCAC	GCCGCCCACC	CCGTCAACGG	CCCCGTCGAC
1234567890	T L E	V G H	A A H L	V N G	P V D
60	Asci	CGCAGGCCGT	CAAGTCCATC	TCGATCGCCG	CACGGATCAT
1234567890	CCGGGCGCGC P G A P	Q A V	K S I	S I A A	R I I
50	ĊCGGCAGGCG	GGCGTGACGC	CCCTGCGCAG	CCACGGACTG	GGGTCCGCG
1234567890	R Q A	G V T P	L R S	H G L	
40	AGGCCGTCGT	дсадсассас	CGCGTCGTCA	TCGAGAACCT	TGAGGCCGAC
1234567890	A V V	О н н	R V V T	E N L	E A D
30	TGGTCGCTGG	CGAAGGTCTG	CGACGCCGCC	CGGCAGCGCA	CTCAGGCGTG
1234567890	WSLE	K V W	D A A	R Q R I	Q A C
10 20 30 40 50 60 70 80 90 100	CTATGTCGAC	GTTTCGCTGG	TCACCCTCGA	GGAAGCCACC	CAAGAGCTCG
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	Y V D	V S L A	T L D	E A T	Q E L A
10	CGCTCTCCCG	CGCTGTCATG	GCCGGTGCCC	CCCTCAGCGA	CACCCAGAIC
1234567890		A V M	A G A L	L S E	T Q I

FIG. 31JJ

CCTGGCGCTA CCGCATCGAC TGGAAGCGCC TCGCGGTCGC 18500 W R R L A V A

77/164

10 34567890	20 1234567890	10 20 30 1234567890 1234567890 123	40 1234567890	50 1234567890	60 1234567890	40 50 60 70 80 90 100 4567890 1234567890 1234567890 1234567890 1234567890 1234567890	80 1234567890	90 1234567890	100 1234567890	
CGAGAGCG P	AACTCGCCGA L A E	GGTCCTCGCC	ESPI BPU1102I GGGCTCAGCC CG	2I CGCGGACACC R T P	Bsu36I ▼ TGAGGTGCCG E V P	TTCTTCTCGA F F S T	TTCTTCTCGA CACTCGAAGG CGCCTGGATC ACCGAGCCGG	CGCCTGGATC A W I	EBP1 Bpull021 BBul361 T ATCGAGAGCG AACTCGCCGA GGTCCTCGCC GGGCTCACC TGAGGTGCCG TTCTTCTCGA CACTCGAAGG CGCCTGGATC ACCGAGCCGG 18100 I E S E L A E V L A G L S P R T P E V P F F S T L E G A W I T E P V	.8100
SCTCGACGG C	Av CACCTACTGG T Y W	KpnI Acc651 TACCTACTGG TACCGCAACC TYN YRN L	TCCGCCACCG R H R	CGTCGGCTTC V G F	GCCCCCGCCG A P A V	GCCCCGCCG TCGAGACCCT A P A V E T L	CGCCACCGAC A T D	GAAGGCTTCA E G F T	KpnI Acc65I ▼ ▼ TGCTCGACGG CACCTACTGG TACCGCACCG CCTCGGCTTC GCCCCCGCG TCGAGACCCT CGCCACCGAC GAAGGCTTCA CCCACTTCAT 18200 L D G T Y W Y R N L R H R V G F A P A V E T L A T D E G F T H F I	.8200
anggtcagc 3 V S	GAGGTCAGC GCCCACCCG TO E V S A H P V	TCCTCACCAT L T M	GACCCTCCCC T L P	GAGACCGTCA E T V T	GAGACCGTCA CCGGCCTCGG E T V T G L G	CACCCTCCGC T L R	CGCGAACAGG GAGGCCAGGA R E Q G G Q E	GAGGCCAGGA G Q E	CGAGGICAGC GCCACCCC TCCTCACCAT GACCCTCCC GAGACCGTCA CCGGCCTCGG CACCCTCCGC CGCGAACAGG GAGGCCAGGA GCGTCTGGTC 18300 E V S A H P V L T M T L P E T V T G L G T L R R B Q G G Q E R L V	18300
S L A	ACCTCACTCG CCGAAGCCTG GACCAACGG	GACCAACGGC T N G	CTCACCATCG L T I D	ATCG ACTGGGGGC	CGICCICCCC	ACCGCAACCG GCCACCACCC T A T G H H P	GCCACCACCC H H P	CGAGCTCCCC E L P	ACCTCACTCG CCGAAGCCTG GACCAACGGC CTCACCATCG ACTGGGCGCC CGTCCTCCCC ACCGCAACCG GCCACCACCC CGAGCTCCCC ACCTACGCCT 18400 T S L A E A W T N G L T I D W A P V L P T A T G H H P E L P T Y A F	18400

pikPKS Sequence

FIG. 31KK

TCCAGCGCCG TCACTACTGG CTCCACGACT CCCCGCCGT CCAGGGCTCC GTGCAGGACT Q R R H Y W L H D S P A V Q G S V Q D S

dneuce
SS
pikPK

		78,	/164	
	8600	8700	8800	8900
1234567890	MIUI	GCCGGCGCCG ACCCCGTACA GCTGGACGTGG TCCCCGCTGG GCGACGCTGG GCGAGGCCCT GGCGGCGGC GGTGGAGCCG 18700	TCGACGGCGT CCTCTCGCTG CTCGCGTGGG ACGAGGGGC GCACCCCGGCC CCTTCACCCG GGGCACCGGC GCCACCCTCA CCCTGGTGCA 18800	GGCGCTGGAG GACGCCCGCCCCCCCCCCCTGGTGC GTGACCCACG GCCCGACCCACG GCCGACCCACG TCACCTCCCC CGCCCAGGCC 18900
	CGACCCCCCC GAGCGCCCCCC GGCTGTCCGCTC GTCGTCGTCC CCGAGGACCG TTCCGCCCCCG GCCGCCCCCG TGCTCGCCGC GCTGTCCGGC 18600 D A S E R A G L S G R W L V V V P E D R S A E A A P V L A A L S G	A G A D P V Q L D V S P L G D R Q R L A A T L G E A L A A A G G A V	D G V L S L L A W D E S A H P G H P A P F T R G T G A T L T L V Q	A L E D A G V A A P L W C V T H G A V S V G R A D H V T S P A Q A
90	TGCTCGCCGC	GGCGGCGCC	GCCACCCTCA	TCACCTCCCC
1234567890	L A A	A A A	A T L T	T S P
80	GCCGCCCGG IGCTCGCCGC	GCGAGGCCT GGCGGCGC GGTGGAGCCG	GGGCACCGGC	GCCGACCACG 1
1234567890	A A P V L A A	E A L A A A G G A V	G T G	A D H V
70 1234567890	TTCCGCCGAG S A E Z	GCGACGCTGG (A T L G	CCTTCACCCG	CGTCGGCCGG
60	CCGAGGACCG	GCGGCTCGCC	CACCCGCCC CC	GCGCGGTGTC
1234567890	B D R	R L A	H P A P	A V S
50 1234567890	GICGICGICC V V P	GCGACCGGCA (D R Q	GCACCCCGGC H P G	GTGACCCACG
40	GCGCTGGCTC	TCCCCGCTGG	ACGAGAGCGC	GCTGTGGTGC
1234567890	R W L	S P L G	E S A	L W C
30 1234567890	GGCTGTCCGG	GCTGGACGTG L D V	SCTG CTCGCGTGGG	TCGCCGCCC A A P
10 20 30 40 50 60 70 80 90 100	GAGCGCGCG GGCTGTCCGG	ACCCCGTACA	CCTCTCGCTG	GACGCCGCC T
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	ERAGLSG	P V Q	L S L	D A G V
1234567890	MluI CGACGCGTCC D A S	GCCGGCGCG ACCCCGTACA GCTGGACGTG TCCCCGC A G A D P V Q L D V S P L	TCGACGGCGT (GGCGCTGGAG A L E

FIG. 31LL

ATGGTGTGG GCATGGCCCG GGTCGCCGCC CTGGAGCACC CCGAGCGGCTG GGCGGCCTTGACCTGC CCTCGGACGC CGACCGGGCG GCCCTGGACC 19000 M V W G M G R V A A L E H P E R W G G L I D L P S D A D R A A L D R

			79/16	64	
	19100	19200		19400	19500
100 1234567890	GCATGACCAC GGTCCTCGC GGCGGTACCGCGGTACGCGCCTCCTCCG GGCTGCTCGC CCGCCGCCTC GTCCGCGCCT CCCTCCCGGC 19100 M T T V L A G G T G E D Q V A V R A S G L L A R R L V R A S L P A	NOLI ©CACGGCACG GCTTCGCCGT GGTGGCACG GTGCTCGTCA CCGGTGCCGA GGAGCCTGCG GCCGCCGAGG CCGCACGCCG GCTGGCCCGC 19200 H G T A S P W W Q A D G T V L V T G A· E B P A A B A A R R L A R	GACGCCCCG GACACCTCCT CCTCCACACC ACCCCCTCCG GCAGCGAAAGG CGCCGAAAGGC ACCTCCGGCCGG CTCCGGCCTC GCCGGGCTCG 19300 D G A G H L L L H T T P S G S E G A E G T S G A A E D S G L A G L V	NOTI TUGCUGAACT CGUGGACUTG GGUGGACGTCGT GAUCTGGCGAC CTUAUGGACG CGGAGGCGGCTG CTUGCCGGCG TUTUCGAUGU 19400 A E L A D L G A T A T V V T C D L T D A E A A R L L A G V S D A	ESPI BDull02I V GCACCGCTC AGCGCCGTCC TCCACCTGCC GACTCGCACG GACGCGCTCG CCGTGTCGT GACCGCGAAG 19500 H P L S A V L H L P P T V D S E P L A A T D A D A L A R V V T A K
40 80 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890	GGTCGCGGTTCCG GGCTGCTCGC CCGCCGCCTC GTCCGCGCCT (Not! ▼ GAGCCTGCG GCCGCCGACGCCG GCTGGCC EPAAR BARRLA	CTCCGGCCTC S G L	NotI ▼ CGGAGGCGC CGCCCGGCTG CTCGCCGGCG E A A A R L L A G V	GACTCCGAGC GGCCGCCGC GACGCGCTCG CCCGTGTCGT GACCGCC
80 1234567890	CCGCCGCCTC R R L	Noti 7 G GCCGCCGAGG A A B A	CCGCCGAGGA A E D	CGCCCGGCTG ARL	GACGCGCTCG D A L A
70 1234567890	GGCTGCTCGC L L A	N ▼ GGAGCCTGCG E P A	ACCTCCGGTG C T S G A	NotI ▼ CGGAGGCGGC E A A	GACCGACGCG T D A
60 1234567890	CGCGCCTCCG R A S G	CCGGTGCCGA G A· E	CGCCGAAGGC A E G	CTCACGGACG L T D A	CGCTCGCCGC
50 1234567890	GGTCGCGGTA V A V	GTGCTCGTCA V L V T	GCAGCGAAGG S E G	CGT GACCTGCGAC CTCACGGACG	GACTCCGAGC D S E P
40 1234567890	GTGAGGACCA E D Q	CGACGGCACG D G T	ACCCCTCCG T P S G	CCACCGTCGT T V V	GCCCACCGIC P I V
30 1234567890	GETCCTCGCC GGCGGTACGG	GGTGGCAGGC W Q A	CCTCCACACC L H T	GGCGCGACGG G A T A	TCCACCTGCC H L P
10 20 30 1234567890 1234567890 12345678	GGTCCTCGCC V L A	G GCTTCGCCGT GGTGGCAGGC CGACGGCA	GACGCCCCG GACACCTCCT CCTCCACACC D G A G H L L L L H T	ACT CGCGGACCTG L A D L	EspI Bpull02I W CACCCGCTC AGGCCGTCC TCCACCTGCC GCCCACCC H P L S A V L H L P P T '
1234567890	GCATGACCAC M T T	GCACGCCACG H G T	GACGGCGCG D G A G	TCGCCGAACT A E L	Espi Bpul W GCACCCGCTC A

ı		0	, o	0	0
	1960	1970	1980	1990	2000
100 1234567890	GCCACCGCCG CGCTCCACCT GGACCGCCTC CTGCGGGGGGCTGC CGGCCCGTCC TGGTCCTCTT CTCCTCGGTC GCCGCGATCT 19600 A T A A L H L D R L L R E A A A G G R P P V L V L F S S V A A I W	GGGGCGGC CGGTANGGCC CCGGTACGGC CTTCCTCGAC GCCCTCGCCG GTCAGCACCG GGCCGACGGC CCCACCGTGA CCTCGGTGGC 19700 G G A G Q G A Y A A G T A F L D A L A G Q H R A D G P T V T S V A	CTGGAGCCCC TGGGAGGGCA CCGGGGGGGGGGGGGGGG	GCCCTGGACA CCGCGCTCGG CCACGGCGAC ACCGCCGCA CGACGACTTCG CCCCCGGCTT CACCACGGCC CGGCCGGGCA 19900 A L D T .A L G H G D T A V T I A D V D W S S F A P G F T T A R P G T	Asci CCCTCCTCGC CGATCTGCCC GAGGGCTCGA CGAGCAGCAG TCGACGACGA CGCCGACGA CACCGTCCTG AGCCGCAGGC TCGGTGCGCT 20000 L L A D L P E A R R A L D E Q Q S T T A A D D T V L S R E L G A L
90	CTCCTCGGTC	CCCACCGTGA	CCCCGCGAC	CACCACGGCC	AGCCGCGAGC
1234567890	S S V	P T V T	P A T	T T A	S R E L
80	TGGTCCTCTT	GGCCGACGGC	CGCCCCTCG	CCCCCGGCTT	CACCGTCCTG
	V L F	A D G	R P L A	P G F	T V L
10 20 30 40 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	CCGCCCGTCC	GTCAGCACCG Q H R	ccrceeccre r e r	TCGAGCTTCG S S F A	CCGCCGACGA A D D
60	CGGAGGCCGT	GCCCTCGCCG	GGCTGCGCCG	CGTCGACTGG	TCGACGACGG
1234567890	G G R	A L A G	L R R	V D W	S T T A
50	CCGCGGCTGC	CTTCCTCGAC	ACCGGGGGAGC	CGATCGCCGA	CGAGCAGCAG
1234567890	A A A	F L D	T G E R	I A D	E Q Q
40	ctgcgggagg	CCGGTACGGC	CGAGGGTGCG	ACCGCCGTCA	GCGCGCTCGA
1234567890	L R E A	G T A	E G A	T A V T	A L D
30	GGACCGCCTC	GCGTACGCCG	GCCGCGTCAC	CCACGGCGAC	Asci GAGGCGCCC E A R R
1234567890	D R L	A Y A A	R V T	H G D	
20	CGCTCCACCT L H L	CGGTCAGGGC G Q G	TGGGAGGGCA W E G S	CCGCGCTCGG	CGATCTGCCC D L P
1234567890	GCCACCGCCG	366666666	CTGGAGCCCC	GCCTGGACA	ccrccrcgc
	A T A A	6 6 A	W S P	A L D T	L L A

pikPKS Sequence

FIG. 31NN

1 1	_		,	0	0
	20100	20200	20300	2040(2050(
100	CACGGGGCGG T G R	regretres.	GCGGGGTCGA G V D	CGAGGACGCG E D A	GGTGGCTTCC
90 1234567890	AGGCCGTCGA (A V D	CCCGGCCACT (CCGGTGGACG (P V D G	TGGCCGGCGG A G G	CTGCCGTGCC
10 20 30 40 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	CACCGGCGCC GAACAGCAGC GCCGTATGCA GGAGTTGGTC CGCGAGCCGTGGT CCTCAACCAC CCCTCCCCG AGGCCGTCGA CACGGGGGG 20100 T G A E Q Q R R M Q E L V R E H L A V L N H P S P B A V D T G R	BBLXI $lacktrigg$ GCCTTCCGTG ACGCGGTCG AGCTCCGCAA CCGCCTCAAG AACGCCACCG GCCTGGCCCT CCCGGCCACT CTGGTCTTCG 20200 A F R D L G F D S L T A V E L R N R L \cdot K N A T G L A L P A T L V F D	ACTACCCGAC CCCCCGGACG TGCTCCTCGC GGAGATCCTG GGCGAGGACG CCGGTGCTGCTT CCGGTGGACG GCGGGTCGA 20300 Y P T P R T L A E F L L A E I L G E Q A G A G E Q L P V D G G V D	CGACGAGCCC GTCGCGATGGC GTGCCGCCTG CCGGGCGTG TCGCCTCGCC	MluI ** ATCTCCGGCT TCCCGCAGGA CCGCGGACGTGGAGG GGCTGTACGA CCCGGACCCTG GACGCGTCCG GGCGGACGTA CTGCCGTGCC GGTGGCTTCC 20500 ISGFPODRGWDVEGLYDPDPDASGRTYCC
70 1234567890	CCTCAACCAC L N H	AACGCCACCG N A T G	CCGGTGCCGG	GGAGGACCTG E D L	MluI GACGCGTCCG D A S G
60	TCGCCGTGGT	CCGCCTCAAG	GGCGAGCAGG	TCGCCTCGCC	CCCGGACCCG
1234567890	A V V	R L ·K	G E Q A	A S P	
50	CGCGAGCACC	AGCTCCGCAA	GGAGATCCTG	CCGGGCGGTG	GGCTGTACGA
	R E H L	L R N	E I L	V D D G	L Y D
40	GGAGTTGGTC	ACGGCGGTCG	TCCTCCTCGC	GTGCCGCCTG	GACGTGGAGG
	E L V	T A V E	L L A	C R L	D V E G
30	GCCGTATGCA	CGACTCGCTG	CTGGCGGAGT	TCGGCATGGC	CCGCGGCTGG
	R M Q	D S L	L A E F	G M A	R G W
20	laccggcgc gaacagc gccgtatgca ggagti	GCCTTCCGTG ACCTCGGATT	CCCCCGGACG	GACGAGCCC GTCGCGATGGC	ATCTCCGGCT TCCCGCAGGA CCGCGGCTGG GACGT
1234567890	t g a e o o r r m o e l	A F R D L G F	PRT	D E P V, A I V G M A	
1234567890	CACCGGCGCC	GCCTTCCGTG	ACTACCCGAC	CGACGAGCCC	ATCTCCGGCT
	T G A	A F R D	Y P T	D E P	I S G F

82/	1	64
-----	---	----

100 1234567890 1234567				82/164		
1034567890 1234567890		20600	20700	20800	3 20900	21000
TOGACCACGC GASCCAACTTC GACCCCAACT TCTTCGGGAT CTCGCCGCGC GAGGCCCTCG CCATGGACCC GCAGCAGCGC CTCCTCCTGC D E A G E P D A D F F G I S P R E A L A M D P Q Q R L L L E GAAGGCCGTC GAGGACGTC GACCTCCTT CAGGGCAGC AGGTCGCGG CTTCGCGGGC ACCAACGGC CTCCTCTGG E A V E D A G I D P I S L Q G Q V G V F A G I N G P H Y E CCGACACACGC CCGACGATCT CTGGGGACGC CAGCATCATC TCGCGGGC CTCCTCTGGA E A V E D A G I D P I S L Q G Q V G V F A G I N G P H Y E CCGACACACGC CGAGGATCT TGAGGGTTAC GTCGGGACGC CAGCATCATG TCGGGGCCGT TCTCGTACACC CCGACACACGC CGAGGATCT TGAGGGTTAC GTCGGGACGCG CAGCATCATG TCGGGCCGT TCTCGTACACC CCGTCACACGG CCGAGGATCT TGAGGGTTAC GTCGGGACGCG CAGCATCATG TCGGGCCGT TCTCGTACACC CCGTCACACGGT CGACGACGTTAC GTCGGGACGC CTGCACCTCT GCGGCCGT TCTCGTACACC CTCGGGCCTG R N I A E D L E G Y V G I G N A A S I M S G R V S Y I L G L A L CCGTCACACGGT CGACGACGTC CGCTGGACTCT CGCGTGCAAG GCCTGCGCAA GGGCCGTG CTTCGTACACC V I V D I A C S S S L V A L H L A V Q A L R K G E C G L A L TOTALACCACC ACCACAACGTT CGTCGACATTC GCCGGCCACC GGAGGACGGC CGGAGAACGGC CGTCGCCCACACACGCC CTTCGCCCACACACGGC CTTCGCCCACACACGCC CTTCGCCCACACAACGGC CCTTCGCCCACACACGC CCTTCGCCCACACACGC CCTTCGCCCACACACCGC CTTCGCCCACACACCGC CCTTCGCCCACACACCGC CCTTCCCCCTCCCCCTCCCCCTCCCCCTCCCCCTCCCCCTCCCC	1234567890	AGACCTCCTG T S W	eccecrecro P L L	GAGGCCCGG	TCGCGGGCGC	GTCGGCGGA(S A D
10 20 30 40 50 60 70 80 1234567890 123457890 123457890 1234567890 12345	90	CICCICCIGG L L L E	CCCACTACGA H Y E	CCTCGGCCTG L G L	BsmI ▼ GGACTGGCGC G L A L	XmnI ▼ CGTTCGCCGC F A A
1034567890 123457890 123457890	80	GCAGCAGCGG	ACCAACGGCC	TCTCGTACAC	GGGCGAATGC	CGGTCGAAGG
	1234567890	Q Q R	T N G P	S Y T	G B C	R S K A
10 20 30 40 50 60 60 60 123456789	70	CCATGGACCC	GTTCGCGGGC	TCGGGCCGTG	CCCTGCGCAA	GGAGGACGGC
	1234567890	M D P	F A G	S G R V	L R K	E D G
TOGACGACGC GGGCGAGTTC GACGCCGACT TCTTCGGGAT CTCGCCGCGCC D E A G E F D A D F F G I S P R GGAGGCCGTC GAGGACGCC GGATCGACCT TCTTCGGGAT CTCGCCGCGCC E A V E D A G I D P T S L Q G Q Q R N T A E D L E G Y V G T G N A A CCGTCACGGT CGACGCCC TGCTCCTCCT CCTGCACCTC V T V D T A C S S S L V A L H L TGTGACGGTC ATGTCGACGC CCACGACGTTC CGCGCAGCCTC V T V B T A C S S S L V A L H L TGTGACGGTC ATGTCGACGC CCACGACGTT CGTGGAGTTC AGCCGGCAGC V T V N S T P T F V E F S R Q R	60	GAGGCCCTCG	AGGTCGGCGT	CAGCATCATG	GCCGTGCAGG	GCGGGCTCGC
	1234567890	E A L A	V G V	S I M	A V Q A	G L A
1234567890 1234567890	50	CTCGCCGCGC	CAGGGGCAGC	GCAACGCCGC	CCTGCACCTC	AGCCGGCAGC
	1234567890	S P R	Q Q Q Q	N A A	L H L	S R Q R
1234567890 1234567890 1234567890 TCGACGAGGC GGGCGAGTTC GACGCCGACT DEAGGCCGTC GAGGACGCCGGACT GGAGGCCGTC GAGGACGCCG GGATCGACCC EAGGCCGTC GAGGACGCCG GGATCGACCC EAGGCCGTC GAGGACGCTC TGAGGGTTAC RNTAR EDLE GY CGTCACGGT CGACACCGC TGCTCCTCT VTV DTAR CGCCCCTCT TGTGACGGTC ATGTCGACGC CCACGACGTT VTV NSTP TF	40	TCTTCGGGAT	GACCTCCCTT	GTCGGGACGG	CGCTGGTCGC	CGTGGAGTTC
	1234567890	F G I	T S L	V G T G	L V A	V E F
1234567890 1234567890 1234567890 1234567890 TCGACGAGGC GGGCGAGTTC D E A G E F GGAGGCCGTC GAGGACGCG E A V E D A G E N T A E D L CGCAACACGG CCGAGGATCT R N T A E D L CGTCACGGT CGACGCCC V T V D T A TGTGACGGTC ATGTCGACGC V T V D T A	30	GACGCCGACT	GGATCGACCC	TGAGGGTTAC	TGCTCCTCCT	CCACGACGIT
	1234567890	D A D F	I D P	E G Y	C S S S	I I F
1234567890 TCGACGAGGC D E A GGAGGCCGTC E A V E A V CGCAACACGG R N T A V T V V T V V T V	20	GGGCGAGTTC	GAGGACGCCG	CCGAGGATCT	CGACACCGCC	ATGTCGACGC
	1234567890	G B F	E D A G	E D L	D T A	M S T P
	10	TCGACGAGGC	GGAGGCCGTC	CGCAACACCG	CCGTCACGGT	TGTGACGGTC
	1234567890	D E A	B A V	R N T A	V T V	V T V

FIG. 31PP

Ω Σ

4 团

H ы

M >

01	20	30	40	20	9	70	80	90	100	
1234567890	1234567890	1234567890 1234567890 1234567890 12345	1234567890	1234567890	1234567890	67890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	
Sacrificado	CGGCGGAGGG	CGTCGGCATG	sphi ▼ crccrcGrcG	AGCGCCTGTC	GGACGCCGC	CGCAACGGAC	ACCGTGTGCT	GGCGGTCGTG	Sphi A CACHTOGACC CGGCGGAGGG CGTCGGCATG CTCCTCGTCG AGCGCCTGTC GGACGCCCGC CGCAACGGAC ACCGTGTGCT GGCGGTCGTG CGCGGCAGCG 21100	1100
а В	о м	D D	л >	പ പ ഗ	Д К	GFGPAEGVGMILVERLSDARRNGHRVLAVVRGSA	고 > 도	V V	ୟ ଓ ଓ	
CGGTCAACCA V N Q	GGACGCGCG D G A	SGTCAACCA GGACGGCGC T V N Q D G A S N G L	TGACCGCCC T A P	GAACGGGCCC N G P	TCGCAGCAGC S Q Q R	GAACGGGCCC TCGCAGCAGC GCGTCATCCG GCGCGCCC GACTGACGA N G P S Q Q R V I R R A L A D A R L T T	GCGCGCGCTC R A L	GCGGACGCCC A D A R	CGGTCAACCA GGACGGCGG AGCAACGGCCCC CAACGGGCCC TCGCAGCAGC GCGTCATCCG GCGCGCGCTC GCGGACGCCC GACTGACGAC 21200 V N Q D G A S N G L T A P N G P S Q Q R V I R R A L A D A R L T T	21200
CGCCGACGTG A D V	GACGTCGTCG D V V E	GCCGACGTG GACGTCGTCG AGGCCCACGG CACGG A D V D V V E A H G T G	CACGGGCACG T G T	CGACTCGGCG R L G D	ACCCGATCGA P I E	CGACTCGGCG ACCCGATCGA GGCACAGGCC CTCATCGCCA CCTACGGCCA GGGGCGCG3 R L G D P I E A Q A L I A T Y G Q G R D	CTCATCGCCA L I A T	CCTACGGCCA Y G Q	CGCCGACGTG GACGTCGTCG AGGCCCACGG CGACTCGGCG ACCCGATCGA GGCACAGGCC CTCATCGCCA CCTACGGCCA GGGGCGCAC 21300 A D V D V V E A H G T G T R L G D P I E A Q A L I A T Y G Q G R D	21300
ACCGAACAGC T E Q P	cecrececer	ACCGAACAGC CGCTGCGCCT GGGGTCGTTG AAGTC T E Q P L R L G S L K S	AAGTCCAACA K S N I	TCGGACACAC G H T	CCAGGCCGCC	PACA TOGGACACO CCAGGCCGCC GCCGGTGTT CCGGCATCAT CAAGATGGTC CAGGCGATGC N I G H T Q A A A G V S G I I K M V Q A M R	CCGGCATCAT G I I	CAAGATGGTC K M V	ACCGAACAGC CGCTGCGCCT GGGGTCGTTG AAGTCCAACA TCGGACACAC CCAGGCGCG GCCGGTGTT CCGGCATCAT CAAGATGGTC CAGGCGATGC 21400 T E Q P L R L G S L K S N I G H T Q A A A G V S G I I K M V Q A M R	21400

pikPKS Sequence

GICGGACCAG AICGACTGGI CGGCGGGCAC GGICGAGCTG CICACCGAGG CCAIGGACTG 21500 Q æ ß œ Ω ø Ω ល

TGGACCGGCC ĸ Ω

CCTGCCGAAG ACGCTCCACG

GCCACGGCGT

U

I ы ۲

> × Д Н

PmlI

84/164

			017.0	•	
	21600.	21700	21800	21900	22000
1234567890	CCCGGTCGAC	GGACGCCTCG	GGGCCGTCGC	GTTCGTGTTC	GCGCTCGCTC
	P V D	G R L A	A V A	F V F	A L A P
90	TCGAAGAAGC	CGCCCAGATC	TTCGAGCACC	GTCGAGTGGC	GTGCGAGGCC
1234567890	E E A	A Q I	F E H R	R V A	C E A
10 20 30 40 50 60 70 80 90 100 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	GCCGAGGAAG CAGGAGGGGG GGCTGCGGCGGTC TCCTCCTTCG GCATCAGCGG CACGAACGTGC TCGAAGAAGC CCCGGTCGAC 21600 PRKQBG G LRRA AVSSPGIS G TNAHIVLEBAPVD	GAGGACGCCC CGGCGGACGA GCCGTCGGTC GGCGTGGCT CGTGTCCGCG AAGACTCCGG CCGCGCTGGA CGCCCAGATC GGACGCCTCG 21700 E D A P A D E P S V G G V V P W L V S A K T P A A L D A Q I G R L A	CCGCGTTCGC CTCGCAGGGC CGTACGGACG CCGCCGATCC GGCCGTAC TGGCCGGCGG GCGTGCGCAG TTCGAGCACC GGGCCGTCGC 21800 A F A S Q G R T D A A D P G A V A R V L A G G R A Q F E H R A V A	Not: W GCTCGGCACC GGACAGGACG ACCTGGCGC CGCACTGGCC GCCTCGGTCGCC TCCGGTGTGG GTCGAGTGGC GTTCGTGTTC 21900 L G T G. Q D D L A A A L A A P E G L V R G V A S G V G R V A F V F	XmmI ▼ CCGGGACAGG GCACGCAGTG GGTGCCGAAC TCCTCGACGT GTCGAAGGAG TTCGCGGCGG CCATGGCCGA GTGCGAGGCC GCGCTC 22000 P G Q G T Q W A G M G A E L L D V S K E F A A M A E C E A A L A P
70 1234567890	CACGAACGCG T N A	AAGACTCCGG K T P A	TGGCCGGCGG A G G	GGGTGTGGCC G V A	1I TTCGCGGCGG F A A A
60 1234567890	GCATCAGCGG I S G	CGTGTCCGCG V S · A	GCTCGCGTAC A R V L	I GTCTGGTCCG L V R	XmnI GTCGAAGGAG 1 S K E E
50	TCCTCCTTCG	reteg tecceteget	GGGCGCGGTC	BBu36I GCGCCTGAGG A P E G	TCCTCGACGT
1234567890	S S F G	V V P W L	G A V		L D V
40 1234567890	CGCGGCCGTC A A V	G G V V	CCGCCGATCC A D P	I CGCACTGGCC A L A	GGTGCCGAAC G A E L
30	GGCTGCGCG	GCCGTCGGTC	CGTACGGACG	NOTI GGACAGGACG ACCTGGCGC CGCACT G. Q D D L A A A L	GGCCGGGATG
1234567890	L R R	P S V	R T D A		A G M
20	CCGAGGAAG CAGGAGGGCG GGCTGCGCCG	gaggacgcc cggcggacga	CTCGCAGGGC	GGACAGGACG	GCACGCAGTG
1234567890	PRKQBGGAGGGCG LRR	E D A P A D E	S Q G	G. Q D D	T Q W
1234567890	GCCGAGGAAG	GAGGACGCCC	CCGCGTTCGC	GCTCGGCACC	CCGGGACAGG
	P R K	E D A P	A F A	L G T	P G G G

FIG. 31RR

	100	5200	2300	2400
1234567890	ABCI CGTACGTGGA CTGGTCGCTG GAGGCCGTCG TCCGACAGGC CCCACGCTGG AGCGGGTCGA TGTCGTCCAG CCCGTGACGT TCGCCGTCAT 22100 Y V D W S L E A V V R Q A P T L E R V D V V Q P V T F A V M	GGTCTCGCTG GCGAAGGTCT GGCGTGACC CCGCAAGCCG TCGTCGGCCA CTCGCAGGGC GAGATCGCCG CCGCGTACGT CGCCGGTGCC 22200	Espl Sphi Bpull021 T CTGAGCCTCG ACGACGCCGC TCGTGTCGTG ACCTGCGCCA GCAAGTCCAT CGGCGCCCAC CTCGCGGGCCGCCAT GCTGTCCTC GCGCTGAGCG 22300 L S L D D A A R V V T L R S K S I G A H L A G Q G G M L S L A L S E	AGGCGGCCGT TGTGGAGCGA CTGGCCGGGT TCGACGGGCT GTCCGTCGCC GCCGTACCGC CACCGTGGTT TCGGGCGACC CGACCCAGAT 22400 A A V V E R L A G F D G L S V A A V N G P T A T V V S G D P T Q I
10 20 30 40 50 60 70 80 90 100 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	CCCGTGACGT	GGTGACC CCGCAAGCCG TCGTCGCCA CTCGCAGGGC GAGATCGCCG CCGCGTACGT CGCCGGTG	EBPI Sphi Bpull02 CGGCGCCCAC CTCGCGGGCC AGGGCGGCAT GCTGTCCCTC GCGCTGAGCGGCGA H L A G Q G G M L S L A L S E	CGGGCT GTCCGTCGC GCCGTACCGC CACCGTGGTT TCGGGCGACC CGACCCAGA:
80 1234567890	TGTCGTCCAG V V Q	GAGATCGCCG E I A A	AGGGCGGCAT G G M	CACCGTGGTT T V V
70 1234567890	AGCGGGTCGA R V D	CTCGCAGGGC S Q G	CTCGCGGGCC L A G Q	GGCCTACCGC P T A
60 1234567890	CCCACGCTGG	TCGTCGGCCA V G H	CGGCGCCCAC G A H	GCCGTCAACG A V N G
50 1234567890	ABCI CCCCGGCGCG P G A	CCGCRAGCCG P Q A V	CTGCGCA GCAAGTCCAT L R S K S I	GTCCGTCGCC S V A
40 1234567890	TCCGACAGGC R Q A	CGGGGTGACC G V T	ACCCTGCGCA T L R S	TCGACGGGCT D G L
30	GAGGCCGTCG E A V V	GGCAGCACCA Q H H	TCGTGTCGTG R V V	TGTGGAGCGA CTGGCCGGGT TO
20 1234567890	CTGGTCGCTG W S L	GTCTCGCTG GCGAAGGTCT GGCAGCACCA V S L A K V W Q H H	CTGAGCCTGG ACGACGCCGC L S L D D A A	TGTGGAGCGA V E R
1234567890	CGTACGTGGA Y V D	GGTCTCGCTG V·S·L	CTGAGCCTGG L S L D	AGGCGGCGT A A V

pikPKS Sequence

CATCGAGAGC 22500 (H) CGGGGTCCGC GCACGGATCA TCCCCGTCGA CTACGCCTCC CACAGGCCC ACGTCGAGAC G V R A R I I P V D Y A S H S A H V E T GTGAGGCCGA CCAAGAGCTC GCTCAGGCGT 4 ø 臼

Α,

				00/104	_
		22600	22700	22800	22900
001 06	1234567890	SCCTCGACG A L D G	TCGAGGTCAG E V S	CACCTCCCTC T S L	CAGACCGAGC Q T E R
06	1234567890	CACCGAACCC (TEP)	ACCCACTTCG :	ACCGCCTCAC R L T	CTACGCCTTC Y A F
08	1234567890	GCGCCTGGAT (MBCI BalI T T T T T T T T T T T T T	GGCGGACAGC	GGCCT CTCTCCTCC CACCACGACC ACCTACCCCAC CTACGCCTTC CAGACCGAGC
70	1234567890	ACCTCGAAG T L E G	Maci Bali TGGCCACCGA	CCGTGACAAC R D N	ACCCACCCG T H P D
9	1234567890	CTTCTTCTCC F F S	GTCGAAACCC V E T L	GCACCCTCCG T L R	CACCACGACC T T T
20	1234567890	CCCAGGTCCC Q V P	CGCCCCGGCC A P A	ACCGGACTCG T G L G	CTCTCCTCCC L L P
40	1234567890	CCCCAGACAC P Q T P	GIGIGGGCTT V G F	CGAGACCGTC E T V	GACTGGGCCT D W A S
30	1234567890	GGGGTTGTCC G L S	CTCCGCCATC LRHR	TGGCGCTGCC A L P	CCTCACCGTC L T V
20	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	GAACTCGCCG ACGTCCTGGC GGGGTTGTCC CCCCAGGTCCC CTTCTTCTCC ACCCTCGAAG GCGCCTGGAT CACCGAACCC GCCTCGACG 22600 E L A D V L A G L S P Q T P Q V P F F S T L E G A W I T E P A L D G	KpnI Acc651 ▼ ▼ GCGCTACTG GTACCGCAAC CTCCGCCATC GTGTGGGCTT CGCCCCGGCC GTCGAAACCC TGGCCACCGA CGAAGGCTTC ACCCACTTCG TCGAGGTCAG 22700 G Y W Y R N L R H R V G F A P A V E T L A T D E G F T H F V E V S	CGCCCACCC GTCCTCACCA TGGCGCTGCC CGAGACCGCG GCACCCTCCG CCGTGACAAC GGCGGACAGC ACCGCCTCAC CACCTCCCTC 22800 A H P V L T M A L P E T V T G L G T L R R D N G G Q H R L T T S L	GCCGAGGCCT GGGCCAACGG CCTCACCGTC GACTCCTCCC CACCACGACC ACCCACCCG ATCTGCCCAC CTACGCCTTC CAGACCGAGC 22900 A B A W A N G L T V D W A S L L P T T T H P D L P T Y A F Q T B R
01	1234567890	GAACTCGCCG E L A D	GCGGCTACTG G Y W	CGCCCACCC	GCCGAGGCCT A E A W

pikPKS Sequence

GGCGCCCG TGCCGCTCGC 23000 G A A V A L A CCCGCTGCTC ы CGGCCGAGCA 耳 囶 CCGGTGACAT CACCTCCGCC GGTCTCGGGGG CTCTCCGCCG (IN SOCIAL S Eco47III ▼

Д

Ħ

æ

×

T F G L

CGGGCGTCGT

GCGACGAGTC

GTCGCCCCGA

GCAGCTGTCC

CGGTCCGTGT

œ

ø

Ö

ß

M

Ω

ຜ

4

Ö

ß

Н ø

TCTACGCGCA

23300 CGTCGTGGCG 23200 AGGCGGAGTG ద ø 9099900909 CGTGCTGCCC Œ Q Ы Д > CCCGGAGGAC CTCACCCTCG ACGCGCCGCT ы Ω 团

4

Ω

႕

H

Ö

囟

ы

Ω

ບ

GACCAGGTCG (D Q V G

CCGAGCCGGG

Ö

4

ø

GGTCGAGGAG

GTTGCGATCT

23100

GCCGGGAACG

CCGTGCTGCT

GTGGCCGGCA (V A G T

GGACCACGCG

CCTGGCTGGC

CGTACGCACC (

CCTCTCCCTC

TCACGGGGAG

Н

ഗ

н

ß

Ö

H

G C L L

GGACTCCGAC

A

ø

ø

Ħ

Ω

K

H

O

Д

ы

1234567890

1234567890

1234567890

1234567890

1234567890

1234567890

1234567890

1234567890

GCCGGTGGAC 23400 TGGCCGCCGC CGGCGCCGA CCCGGAGGCC 4 闰

ρι

ø

CCGTGCGGAC (

TGCTGGCCGC

H

GCCACCGGTG 7

GACGCGCCAC

CCGTCGCCGA

23500

GAGGTGTTCG

GCGTGGCGAC

GTGTCTGGCG

GGCGTCCGTG

CITCGCGGCG AACGGCTACG GCTACGGCCC CCTCTTCCAG

ßų

А

O

œ

œ

3

>

Ö

2

>

Ö

œ

ы

Ö

×

Ø

×

Ö

ď

4

ß,

臼

TGTACGAGCG

GTGGACGGTC

Eco47III

pikPKS Sequence

88/164

		· ·	337 101		
	23600	23700	23800	23900	24000
100	CCGGTGCGGG	GGCCCCGCC	GCGCAGCTGG	TGCTGGGCGG	CCGGGGCGAG
1234567890	G A G	A P A	A Q L A		R G E
90	GTGCAGGCGG	GCGTGCGGCT	CGTCGACCCC	GGCGCGGTCG	AGGCCGTCGA
1234567890	V Q A A	V R L	V D P		A V D
80	CGACGCCGCC	ACCGCCCTCC	CGGTGCTGCC	GGCCTGCCC	GACCTGGTGG
1234567890	D A A	T A L R	V L P	A L P	D L V E
70	CGGCGCTGCT	GGTCGGCGCC	GACTCCCTCA	GTGCCGCGCA	GGACCTTACG
1234567890	A L L	V G A	D S L T	A A Q	D L T
60	GCCTTCACC	TCCTGTACGC	GTTCGCCGCG	GCCTGGGACG	TGTCCTTCCC
1234567890		L Y . A	F A A	A W D G	S F P
50	CGCGCGGTTC	GAGCGGGATC	GGCAGCCGGT	GGAGTGGACC	ACCGAGGICC
1234567890	A R F	E R D L	Q P V	E W T	T E V L
40	GTGCCGAGGG	OGTTCGCCTG	GACTCCTCCG	TGCACCTGCT	CGCCGGTGGC
1234567890	A E G	V R L	D S S G	H L L	
30	GAGGTCGCCG	ACGCGGCTGC	GAGCGCCGCC	CTGGACGCGC	CGGCGCTGCG
1234567890	E V A G	A A A	S A A	L D A L	A L R
10 20 30 40 50 60 70 80 90 100	CCGACGTGGC CTGCCGGCCG GTGCCGAGGG CGCGCGGTTC GGCCTTCACC CGGCGCTGCT CGACGCCGCC GTGCAGGCGG CCGGTGCGGG 23600	CCGGGGCGTT CGGCGGGCG ACGCGGCTGC CGTCGGGGATC TCCTGTACGC GGTCGGCGCC ACCGCCCTCC GCGTGCGGCT GGCCCCCCCCCC	GGCCCGGACA CGGTGTCCGT GAGCGCCGCC GGCAGCCGGT GTTCGCCGCG GACTCCCTCA CGGTGCTGCC CGTCGACCCC GCGCAGCTGG 23800	CGGCCTTCAG CGACCCGACT CTGGACGTGCT GGAGTGGACC GCCTGGGACG GTGCCGCGCA GGCCCTGCCC GGCGGGTCG TGCTGGGCGG 23900	CGACGCCGAC GGTCTCGCCG CGCCGCTGCC ACCGAGGTCC TGTCCTTCCC GGACCTTACG GACCTGGTGG AGGCCGTCGA CCGGGGCGAG 24000
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	DVALPAVO A SARGO ARFGLHPALL DAAVQAA GAG		G P D T V S V S A A D S S G Q P V F A A D S L T V L P V D P A Q L A	A F S D P T L D A L H L L B W T A W D G A A Q A L P G A V V L G G	DADGLAAAALRAGGTTES ON SFPDLT DLVEAA
1234567890	CCGACGTGGC	CCGGGGCGTT	GGCCCGGACA	CGGCCTTCAG	CGACGCCGAC
	D V A	R G V	G P D T	A F S	D A D

89/	164
-----	-----

			•		
	24100	24200	24300	24400	24500
100	GCGCTGATGC	GGTCCACGGG	GGCCGGGGAC	CTCGGCTCGG	CGGCCGACGG
1234567890	A L M Q	S T G	A G D	L G S G	A D G
90	TGGGCCGGAG CATGTCCGCG AGGCCCTGCA CGGGTCGCTC GCGCTGATGC	GACGGCCTGC	AAGCCCGGAC	CGCGACCGCC	GGCCTCGCCG
1234567890	G P E H V R E A L H G S L A L M Q	D G L R	A R T	A T A	
80	AGGCCCTGCA	CCGTTCCGGC	CTCGCCGGGG	GCAGCGCCT	CGCGGCCGAC
1234567890	A L H	R S G		S A L	A A D
70	CATGICCGCG	CGGTCGCCGC	CCTGCTCGAC	GCCGCCCTCG	CCGCGCCCGC
1234567890	H V R E	V A A	L L D	A A L G	
60	TGGGCCGGAG	ACCCGCGACG	GCCGGTTCGT	CTCTGGAGAC	GCGCGGGCCG
1234567890	G P E	T R D A	R F V	S G D	A R A A
50	GCCG CCGGCCCCGA	GGTGCTCGTG	GAGAGCCCGG	TCGGCGGCAC	ACCCGCCTG
1234567890	A A G P D	V L V	ESPG	G G T	P R L
40	TGCCCCGCCG	ATGGGCGCCT	CGCGCAGACG	GACGCCACCG	CGCTCCTCGT
1234567890	C P A A	G R L	A Q T	D A T V	L L V
30	CCTGGTGGCC	CGGTTCACCG	TCGGCCGGTC	GACGACCGGG	CGGGACGGGG
	L V A	R F T D	G R S	T T G	R D G A
10 20 30 40 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	ACCCCGGCCC CGGCGACCGT CCTGGTGGCC TGCCCCGCG CGGCCCGGCG AGGCCCTGCA CGGGTCGCTC GCGCTGATGC 24100 TPAPAT YLV LVAC PAAGPOCCGCCCCCCCGATGGCCGGAG CATGTCCGCG AGGCCCTGCA CGGGTCGCTC GCGCTGATGC 24100	AGGCCTGGCT GGCCGACGACG ATGGGCGCCT GGTGCTCGTG ACCCGCGACG CGGTCGCCGC CCGTTCCGGC GACGGCCTGC GGTCCACGGG 24200 A W L A D E R F T D G R L V L V T R D A V A A R S G D G L R S T G	ACAGGCCGCC GTCTGGGCCGGTC CGCGCAGACG GAGAGCCCGG GCCGGTTCGT CCTGCTCGAC CTCGCCGGGG AAGCCCCGGAC GGCCGGGGAC 24300 Q A A V W G L G R S A Q T E S P G R F V L L D L A G E A R T A G D	GCCACCGCCG GGGACGCCTCGGCGGCCACCG TCGGCGCCAC CTCTGGAGAC GCCGCCCTCG GCAGCGCCCT CGCGACCGCC CTCGGCTCGG	GCGAGCCGCA GCTCGCCTC CGGGACGGG CGCTCCTCGT ACCCCGCCTG GCGCGGCCG CCGCGCCCGAC CGCGGCCGAC GGCCTCGCCG CGGCCGACGG 24500 E P Q L A L R D G A L L V P R L A R A A A A A A D G L A A D G
10	ACCCGGCCC	AGGCCTGGCT	ACAGGCCGCC	GCCACCGCCG	GCGAGCCGCA
1234567890	T P A P	A W L	Q A A	A T A G	B P Q

FIG. 31WW

90	/1	64
	, .	•

		_			
	24600	24700	24800	24900	25000
100	CCTCGCCGCT CTGCCGCTGC CGCCCTCTGG CGTCTGGAGC CCGGTACGGA CGGCAGCCTG GAGAGCCTCA CGGCGGCGC CGGCGACGCC 24600	GAGACCTTCG CCCCGGAGCCCG GGACAGGTCC GCATCGCGAT CCGGGCCACC GGTCTCAACT TCCGCGACGT CCTGATCGCC CTCGGCATGT 24700	ACCCCGATCC GGCGCTGATG GGCACCGAGG GAGCCGCGG ACCGGCCCCG GCGTCACGCA CCTCGCCCCC GGCGACCGGG TCATGGGCCT 24800	GCTCTCCGGC GCGTACGCCC CGGTCGTCG CGGACCGTCG CGCGGATGCC CGAGGGGGGTGG ACGTTCGCCC AGGGCGCCTC CGTGCCGGTG 24900	GIGITICCTGA CGCCCTCTA CGCCTGCCG CGCGCGA ACGTCAAGCC CGGCGAGCGC CTCCTGGTCC ACTCCGCCGC CGGTGGCGTG GGCATGGCCG 25000
1234567890	L A A L P L P A A P A L W R L E P G T D G S L E S L T A A P G D A	B T L A P E P L G P G Q V R I A I R A T G L N F R D V L I A L G M Y	P D P A L M G T E G A G V V T A T G P G V T H L A P G D R V M G L	L S G A. Y A P V V V A D A R T V A R M P E G W T F A Q G A S V P V	V F L T A V Y A L R D L A D V K P G E R L L V H S A A G G V G M A A
90	CGGCGCCC	CCTGATCGCC	GGCGACCGGG	AGGGCGCCTC	cecreccere
1234567890	A A P	L I A		G A S	G G V
80	GAGAGCCTCA	TCCGCGACGT	CCTCGCCCCC	ACGITCGCCC	ACTCCGCCGC
1234567890	E S L T	R D V	L A P	I F A Q	S A A
70	CGGCAGCCTG	GGTCTCAACT	GCGTCACGCA	CGAGGGGTGG	CTCCTGGTCC
1234567890	G S L	G L N F	V T H	E G W	L L V H
60	CCGGTACGGA	CCGGGCCACC	ACCGGCCCCG	CGCGGATGCC	CGGCGAGCGC
1234567890	G T D	R A T	T G P G	R M P	G E R
50	CGTCTGGAGC	GCATCGCGAT	GGTCACCGCG	CGGACCGTCG	ACGTCAAGCC
1234567890	R L E P	I A I	V T A	R T V A	V K P
40	GGCCCTCTGG	GGACAGGTCC	GAGCCGGCGT	GGCGGACGCG	GACCTGGCGG
1234567890	A L W	G Q V R	A G V	A D A	D L A D
30	CCGCCGCTCC	GCTCGGCCCG	GCCACCGAGG	CGGTCGTCGT	CGCCCTGCGC
1234567890	A A P	L G P	G T E G	V V V	A L R
10 20 30 40 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	CTCGCCGCT CTGCCGCTGC GGCCC	CCCCGGAGCC	GGCGCTGATG A L M	GGC GCGTACGCCC CGGTCGTCGT	CGGCCGTCTA A V Y
10	CCTCGCCGCT	GAGACCCTCG	ACCCGATCC	GCTCTCCGGC	GIGITCCIGA CC
1234567890	L A A	B T L A	P D P	L S G	V F L T

FIG. 31XX

91	/1	64
----	----	----

			-	_
	25100	25200	25300	25400
1234567890	COGTGCAGCT CGCCCGGCAC TGGGGCGTGG AGGTCACGGCGAGT CACGGGAAGT GGGACGCCCT GCGCGCCTC GGCCTGGACG ACGCGCACAT 25100 V Q L A R H W G V E V H G T A S H G K W D A L R A L G L D D A H I	CGCCTCCTCC CGCACCCTGG ACTTCGAGTC CGCGTTCCGTTC	GACCCCTCGC TGCGCCTGCT CGGGCCGGGTTCG TGGAGATGGG GAAGACCGAC GTCCGCGACG CGGAGCGGGT CGCCGCCGAC CACCCCGGTG 25300 D A S L R L L G P G G R F V E M G K T D V R D A E R V A A D H P G V	TCGGCTACCG CGCCTTCGAC CTGGGCGGAGCCGGA GCGGATCGGC GAGATGCTCG CCGAGGTCAT CGCCCTCTTC GAGGACGGGG TGCTCCGGCA 25400 G Y R A F D L G E A G P E R I G E M L A E V I A L F E D G V L R H
90	GGACGCCCT GCGCGCCTCGACG ACGCGCACA.	CGCTCGCCCG	CGCCGCCGAC	GAGGACGGG
1234567890	D A L R A L G L D D A H I	L A R	A A D	E D G V
80	GCGCGCGCTC	TATGGACGTC GTACTGAACT CGCTCGCCCG	CGGAGCGGGT	CGCCTCTTC
1234567890	R A L	M D V V L N S L A R	E R V	A L F
70	GGGACGCCT	CATGGACGTC	GAAGACCGAC GTCCGCGACG CGGAGCGGGT	CCGAGGTCAT
1234567890	D A L	M D V	K T D V R D A E R V	E V I
60	CACGGGAAGT G	GCGGGGCGGG	GAAGACCGAC	GAGATGCTCG
1234567890	H G K W		K T D	E M L A
40 50 60 70 80 90 100	CACGGCGAGT	GCCGCTTCCG GCGGGGCGGG (A A A S G G A G	TGGAGATGGG	GCGGATCGGC
67890 1234567890 1234567890 1234567890 1234567890	T A S		E M G	R I G
40	AGGTCCACGG	CGCGTTCCGT	GGCCGGTTCG	CCGGGCCGGA
1234567890	V H G	A F R	GRFV	G P E
30	TGGGCGTGG	ACTTCGAGTC	9 d 9	CTGGGCGAGG
1234567890	W G V E	F E S	29992239922	L G E A
10 20 30 1234567890 1234567890 12345	GTGCAGCT CGCCCGGCAC TGGGCGTGG AGG	GCCTCCTCC CGCACCCTGG ACTTCGAGTC CGCGT A S S R T L D F E S A F	GACGCCTCGC TGCGCCTGCT CGGGCCGGGC GGCCG	CGCCTTCGAC A. F. D
1234567890	COGTGCAGCT	CGCCTCCTCC	GACGCCTCGC	TCGGCTACCG
	V Q L	A S S	D A S L	G Y R

FIG. 31YY

CCTGCCCGTC ACGACCTGGG ACGTGCGCCG GGCCCGCGAC GCCTTCCGGC ACGTCAGCCAGCCACA ACGGGCAAGG TCGTCCTCAC GATGCCGTCG 25500 L P V T T W D V R R A R D A F R H V S Q A R H T G K V V L T M P S

SERIAL NO.: 09/988,384

REPL	ACEN	JENT	SHEET
------	-------------	-------------	-------

e)uce
Seque
IKPKS
Ω.

		92/	164	
	55600	25700	25800	25900
100 1234567890	Pmli GGCCTCGACC CGGAGGGTAC GGTCCTGCTG ACCGGCGCA CCGGTGCGCT GGGGGCATC GTGGCCCGGC ACGTGGTGGG CGAGTGGGGC GTACGACGCC 25600 G L D P E G T V L L T G G T G A L G G I V A R H V V G E W G V R R L	Apali ** TGCTGCTCGT GAGCCGGCGG GGCACGACG CCCCCGGGCGCGCCCCT GGGAGCCGAC GTCTCGGTGG CCGCGTGCGA 25700 L L V S R R G T D A P G A G E L V H E L E A L G A D V S V A A C D	CGICGCCGAC CGCGAAGCCC ICACCGGCGIACICGACTCGCGCT CACCGCGGIC GICCACAGGG CAGGCGICCT CICCGACGGC 25800 V A D R E A L I A V L D S I P A E H P L I A V H I A G V L S D G	ACCCTCCCT CGATGACAGGATGTG GAACACGTAC TGCGTCCCAA GGTCGACGCC GCGTTCCTCC TCGACGAACT CACCTCGACG CCCGGCTACG 25900 T L P S M T A E D V E H V L R P K V D A A F L L D B L T S T P G Y D
50 60 70 80 90 100 1234567890 1234567890 1234567890 1234567890	CGAGTGGGGC E W G	Apali ▼ . CGGCGAGCTC GTGCAGGCCCT GGGAGCCGAC GTCTCGGTGG CCGCGTGCG G E L V H E L E A L G A D V S V A A C D	CAGGCGTCCT G V L	RCTCCCAA GGTCGACG GCGTTCCTCC TCGACGAACT CACCTCGACG CCCGGCTACG
80 1234567890	PmlI ACGTGGTGGG V V G	GGAGCCGAC G A D	GTCCACACGG V H T A	TCGACGAACT D E L
70 1234567890	Pmli CGGTGCGCT GGGGGCATC GTGGCCCGGC ACGTGGTGGG	TGGAGGCCCT E A L	ACACCCGCT CACCGCGGCGTCCT HPLTAVVHTAGVL	GCGTTCCTCC A F L L
60 1234567890	GGGGGCATC	Apali ▼ GTGCACGAGC V H E L	AACACCCGCT H P L	GGTCGACGCC V D A
50 1234567890	ccegrececr G A L	CGGCGAGCTC G E L	ATCCCCGCCG AA I P A E	TGCGTCCCAA R P K
1234567890	ACCGGCGCA T G G T	CCCCGGGGGG	ACTCGACTCG L D S	GAACACGTAC 7 E H V L
30 1234567890	GGTCCTGCTG V L L	BAGCCGGCGG GGCACGGACG CCCCGGGCG S R R G T D A P G A	TCACCGCCGT T A V	artgacagc ggaggatgtg gaaca(M T A E D V E H
10 20 30 40 1234567890 1234567890 1234567890	GOCCTCGACC CGGAGGGTAC GGTCCTGCTG ACCGGC	GAGCCGGCGG S R R	Y A D R E A L T A V L D S	CGATGACAGC M T A
10 1234567890	GCCTCGACC	TGCTGCTCGT L L V	CGTCGCCGAC V A D	ACCCTCCCT C

CTTCGGTGGC GCGGCGTACGC CGCCGCCAAC GCCACCCTCG ACGCCCTCGC 26000 F G G A Y A A A N A T L D A L A ACCIGGCAGC GITCGICATG ITCTCCTCCG CCGCCGCCGT

10	20	30	40	50	9	70	80	06	100	
1234567890	12345678	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
CTGGCGCCGC W R R	CTGGCGCCGC CGGACAGCCG GACTCCCCCC CCTCTCGGGGCC TCTGGGCCGA GACCAGCGGC ATGACCGGCG GACTCAGCGA CACCGACCGC 26100 W R R R R T A G L P A L S L G W G L W A E T S G M T G G L S D T D R	GACTCCCGC L P A	ccrcrccrc r s r	GGCTGGGGCC	TCTGGGCCGA W A E	GACCAGCGGC TSG	ATGACCGGCG M T G G	GACTCAGCGA L S D	CACCGACCGC T D R	26100
TCGCGGCTGG S R L A	TCGCGGCTGG CCCGTTCCGG GGCGACGCC ATGGACAGCG AGCTGACCCT GTCCCTCTG GACGCGGCCA TGCGCCGCGA CGACCCGGCG CTCGTCCCGA 26200 S R L A R S G A T P M D S E L T L S L L D A A M R R D D P A L V P I	GGCGACGCCC A T P	ATGGACAGCG M D S E	AGCTGACCCT L T L	GTCCCTCCTG S L L	GACGCGGCCA D A A M	TGCGCCGCGA R R D	CGACCCGGCG D P A	CTCGTCCCGA L V P I	26200
TCGCCCTGGA A L D		CGTCGCCGCG CTCCGCGCCC AGC	AGCAGCGCGA Q R D	Sphi CGGCATGCTG G G M L A	I GCGCCGCTGC A P L L	Espi Bpull02I TCAGCGGGCT S G L	CACCCGCGGA TRG	TCGCGGGTCG S R V G	Asci GCGCGCGCC G A P	26300
GGTCAACCAG V N Q	GGTCAACCAG CGCAGGCAGG CGCGGGCGAG GCGGACACGG ACCTCGGCGG GCGGCTCGCC GCGATGACAC CGGACGACCG GGTCGCGCAC 26400 V N Q R R A A A G G A G E A D T D L G G R L A A M T P D D R V A H	CCGCCGGAGG	CGCGGGCGAG A G E	GCGGACACGG A D T D	ACCTCGGCGG L G G	GCGGCTCGCC R L A	GCGATGACAC A M T P	CGGACGACCG D D R	GGTCGCGCAC V A H	26400
CTGCGGGACC L R D L	Pmli TGCGGGGACC TCGTCCGTAC GCACGTCCTGG GACACGGCAC CCCGAGCCGG GTGGACCTGG AGCGGGCCTT CCGCGACACC GGTTTCGACT 26500 L R D L V R T H V A T V L G H G T P S R V D L E R A F R D T G F D S	PmlI GCACGTGGCG H V A	ACCGTCCTGG T V L G	GACACGGCAC H G T	CCCGAGCCGG PSR	GTGGACCTGG V D L E	AGCGGGCCTT R A F	CCGCGACACC R D T	GGTTTCGACT G F D S	26500

pikPKS Sequence

94/1	64
0	

د 26600	6700	800	000	8
υ	(4	26	269	270
GGGAGCTCG	GCAGACCACG	AGGGCGCTGG	GCGACTCCGA	CTTCTGACCT GCCCGACACC ACCGGCACCA CCGGCACCAC CAGCCCCCT CACACACGGA ACACGGAAGG GACAGGGGAG AACGGGAGCC ATGGCGAACA 27000
E L A	Q T T	R A L A	D S D	R A N N
CCCACCCGG	CGACCGATCG T D R	CGCCCGGCTC A R L	AAGGAGCTGG K E L G	AACGGGAGCC
CTTCGACCAC	ACGGCCTCGG	CGGAGCTCGC	CTTCATCGAC	GACAGGCGAG
F D H	T A S A	E L A	F I D	
CCACGCTGGT	GTCCGGAGAC	GCCGCCGTC	ACCTCTTCTC	ACACGGAACG
T L V	S G D	G G R P	L F S	
CGGCTGCCGG	AAGGCACCGG	GCCCGCCGCC	TCCGACGACG	CACACACGGA
R L P A	G T. G	P A A	S D D D	
GACCGGGCTG T G L	TCCTGGGCGG S W A E	ccrcccrcgc s L A	CXI GGACGAGGCG D E A	CAGCCCCCCT
TCAACGCCGC N A A	CGCGGGCGGG A G G	GGCGTGCTCG	BB BG CCACCGACCT T D L	CCGGCACCAC
CGCAACCGIC	TCGCCACGGC	CCGGCTGGAA	GGCGACGACG	ACCGGCACCA
R N R L	A T A	R L B	G D D A	
CGTCGAACTC	CTCGACGAAC	CCGAACTCGA	GGGGGACGAC	GCCCGACACC
V E L	L D E L	E L D	G D D	
CGCTCACCGC	CGGGCACCTG	GCGGCCCTCG	CCGCGGCCCT	CITCIGACCI
L T A	G H L	A A L A	A A L	
	CGCTCACCGC CGTCGAACTC CGCAACCGTC TCAACGCCGC GACCGGCTGCTGCTGGT CTTCGACCAC CCCACCCCGG GGGAGCTCGC 26600	CGCTCACCGC CGTCGAACTC TCAACGCCGC GACCGGGTTG CGCTGCCGG CCACGCTGGT CTTCGACCAC CCCACCCCGG GGGAGCTCGC 26600 L T A V E L R N R L N A A T G L R L P A T L V F D H P T P G E L A CGGGCACCTG CTCGACGAAC TCGCCACGGC GGCGGCGG TCCTGGGCGG AAGGCACCGG GTCCGGAGGAC ACGGCCTCGG CGACCGATCG GCAGACCACG 26700 G H L L D E L A T A A G G S W A E G T G S G D T A S A T D R Q T T	CGCTCACCGC CGTCGAACTCT CTAACGCCGC GACCGGGCTG CGGCTGCTGGT CTTCGACCAC CCCACCCCGG GGGAGCTCGC 26600 L T A V E L R N R L N A A T G L R L P A T L V F D H P T P G E L A CGGGCACCTG CTCGACGAAC TCGCGGGGGGG TCCTGGGCGG AAGGCACCGG GTCCGGAGAC ACGGCCTCGG CGACCGATCG GCAGACCACG 26700 G H L L D E L A T A A G G S W A E G T G S G D T A S A T D R Q T T GCGGCCTCG CCGAACTCGAA GGCGTGCTCG CCTCCCTCGC GCCGCCGCC GGCGGCCGTC CGGAGCTCGC AGGCGCTGG 26800 A A L A E L D R L E G V L A S L A P A A G G R P E L A A R L R A L A	CGCCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

FIG. 31BBB

95	/1	64
----	----	----

		•		
	SfcI ▼ ▼ B T B D K L R D Y L K R V T A E L Q Q N T R R L R E I E G R T H E P V	GGCGATGGTG GGCATGGCCT GCCTGGCGGTGTC GCCTCGCCCG AGGACCTGTG GCAGCTGGTG GCCGGGGACG GGGACGCGAT CTCGGAGTTC 27200 A I V G M A C R L P G G V A S P E D L W Q L V A G D G D A I S E P	Miui CCGCAGGACC GCGGCTGGGA CGTGGACC CCGACCCGGA CGCGTCCGGC AGGACGTACT GCCGGTCCGG CGGATTCCTG CACGACGCCG 27300 P Q D R G W D V E G L Y D P D P D A S G R T Y C R S G G F L H D A G	GCGAGITCGA CGCCGACTIC TICGGGATCI CGCCGCGGG GGCCCTCGCC AGCAGCGACT GTCCCTCACC ACCGCGTGGG AGGCGATCGA 27400 E F D A D F F G I S P R E A L A M D P Q Q R L S L T T A W E A I E
100	GGT V	TTC	CGGATTCCTG CACGACGCCG G F L H D A G	rcga B
1567	AGCC P	TCGGAG S E	GACC D 1	CGAT
1234	ACG E	CTC S	CAC H	AGG
90	292 -	GAT	ic L	7 8 8
1567	SG &	ACGC A	ATTC	6CG1
1234	GGA(GGG D	තු ල	ACC
890	PBLI SfcI ▼ ▼ CGAGCTGCAGCGTCTGCG CGAGATCGAG GGACGCACGC ACGAGCCGGT CGAGCTGCAG CAGAACACCA GGCGTCTGCG CGAGATCGAG GGACGCACGC ACGAGCCGGT	GCCTCGCCCG AGGACCTGTG GCACGGGGACG GGGACGCGATA	MluI TGACC CCGACCCGGA CGCGTCCGGC AGGACGTACT GCCGGTCCGG	T.
1567	JATC I	විරිදි විරිදි	GGTC	CCTC
1234	CGAC	000 g	000 8	GIC
70	ව ය	GTG V	PACT C	HCT L
1567	Fren L	GCTG	ACGT T	AGCC
123,) 200 200	ପ୍ରି ଫ	AGG R	AGC
890	200 A	GTG *	ည ၁၉၅	ည္သည္
1567	AACA T	PCC.	I GTCC S	GACC D 1
123	Pati I ▼ AG CAG	AGG	MluI CGCG	ATG
890	Pa SfcI ▼ ▼ TGCAG	£000 €	G D	CGCC A
4567	S. ► Geric	TCG(s	PCC P	iccr L
123	කු ස	000 4	80 0	99 ₹
40) A	GGTGTC G V	aACC P	3CGA B
4567	TCACC	ල	TAC	2000
123	909 •	წ ი	CTGT.	99
30	LAGC R	ط م	CTGGAGGGG	ATCT I S
4567	ភ្ជា ភ្ជា		GGA	999
123	TAC	ည္မ	CGT	TTC
890	GAC D	GCATGGCCT GCCGCCTGCC	d D	CTTC
4567	£ 2000 2000	ATGC M 7	GCT(CG.
10 20 30 40 50 60 70 80 90 103 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	GAAGACAA GCTCCGCGAC TACCTCAAGC E D K L R D Y L K R	ည္သမ	CCGCAGGACC GCGGCTGGGA P Q D R G W D	. gg ∢
10	K M	У V	HCC R	rcga D
4567	AAG.	GATCGT	ධ් ය I	AGT
123	ACG	99 P	ر م م	9 9 8

FIG. 31CCC

GAGÇGCGGGC ATCGACCCGA CGGCCCTGAA GGGCAGCGGC CTCGCGCGGC CTGGCACACC GGCTACACCT CGGGGCAGAC CACCGCCGTG 27500 S A G I D P T A L K G S G L G V F V G G W H T G Y T S G Q T T A V

96	/1	64
----	----	----

10	20	30	40	50	09	70	80	06	100	
1234567890	1234567890	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
CAGTCGCCG Q S P E	CAGTCGCCCG AGCTGGAGGG CCACCTGG1 Q S P E L E G H L V	CCACCTGGTC H L V	AGCGGCGCGG S G A A	GCGG CGCTGGGCTCCTCGCGT ACGTCCTCGG TACGGACGGA CCGGCCCTGA	ccrerccesc L S G	CGTATCGCGT R I A Y	ACGTCCTCGG V L G	TACGGACGGA T D G	CAGTCGCCCG AGCTGGAGGG CCCTGGGCGTT CCTGTCCGGC CGTATCGCGT ACGTCCTCGG TACGGACGGA CCGGCCCTGA 27600 Q S P E L E G H L V S G A A L G F L S G R I A Y V L G T D G P A L T	27600
CCGTGGACAC V D T	GGCCTGCTCG A C S	GIGGACAC GGCCIGCTCG ICCTCGCTGG	TCGCCCTGCA A L H	creca ccreecere L H L A V	CAGGCCCTCC Q A L R	CAGGCCCTCC GCAAGGGGGA GTGCGACATG GCCCTCGCCG GTGGTGTCA(Q A L R K G E C D M A L A G G V T	GTGCGACATG C D M	GCCCTCGCCG A L A G	CCGTGGACAC GGCCTGCTCG TCGCCCTGCA CCTCGCCGTG CAGGCCCTCC GCAAGGGCGA GTGCGACATG GCCCTCGCCG GTGGTGTCAC 27700 V D T A C S S L V A L H L A V Q A L R K G E C D M A L A G G V T	27700
GGTCATGCCC V M P	AACGCGGACC N A D L	GTCATGCCC AACGCGGACC TGTTCGTGCA GTTC V M P N A D L F V Q F	GTTCAGCCGG F S R	CAGCGCGGC Q R G L	TGGCCGCGGA A A D	XmnI CAGCGGGGC TGGCCGGGA CGGCCGGTCG AAGGCGTTCG CCACCTCGGC Q R G L A A D G R S K A F A T S A	XmnI AAGGCGTTCG K A F A	CCACCTCGGC T S A	XmmI ▼ ▼ V M P N A D L F V Q F S R G R G L A A D G R S K A F A T S A D G F	27800
GCCCCGCGG	AGGGGCCCGG	GGCCCCGCGG AGTCCTGCTG GTGGA G P A E .G A G V L L V E	GTGGAGCGCC V E R L	R L S D A	CCGCCGCAAC R R N	BamHI GGACACCGGA TC G H R I	BamHI ▼ GGACACCGGA TCCTCGCGGT G H R I L A V	CGTCCGCGCC V R G	BamHI GGCCCCGCGG AGGGCGCCGG AGTCCTGCTG GTGGAGGGCC TGTCGGACGC CCGCCGCAAC GGACACCGGA TCCTCGCGGT CGTCCGCGGC AGCGCGGTCA 27900 G P A E .G A G V L L V E R L S D A R R N G H R I L A V V R G S A V N	27900
ACCAGGACGG Q D G	GCCAGCAAC	CAGGACGG CGCCAGCAAC GGCCTCACGG	CTCCGCACGG	SCACGG GCCTCCCAG CAGCGCGTCA TCCGACGGGC CCTGGCGGAC GCCCGGCTCG CGCCGGGTGA	CAGCGCGTCA Q R V I	TCCGACGGGC R R A	CCTGGCGGAC L A D	GCCCGGCTCG A R L A	ACCAGGACGG CGCCAGCAACGG CTCCGCACGG GCCCTCCCAG CAGCGCGTCA TCCGACGGGC CCTGGCGGAC GCCCGGCTCG CGCCGGGTGA 28000 Q D G A S N G L T A P H G P S Q Q R V I R R A L A D A R L A P G D	28000

FIG. 31DDD

10	20	30	40	50	09	10	80	06	100	
1234567890	12345678	1234567890	1234567890	1234567890	1234567890	67890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	
CGTGGACGTC V D V	CGTGGACGTC GTCGAGGCGC ACGCCACGCGCTC GGCGACCCGA GGCCCTCATC GCCACCTACG GCCAGGAGAA GAGCAGCGAA 28100 V D V V E A H G T G T R L G D P I E A Q A L I A T Y G Q E K S S E	ACGGCACGGG	CACGCGCTC T R L	GGCGACCCGA G D P I	TCGAGGCGCA E A Q	GGCCCTCATC A L I	GCCACCTACG A T Y G	GCCAGGAGAA Q E K	GAGCAGCGAA S S E	28100
CAGCCGCTGA Q P L R	CAGCCGCTGA GGCTGGGCC GTTGAAGTCG AACATCGGGC ACACGCAGGC CGCGGCCGGT GTCGCAGGCG ATGCGCCACG 28200 Q P L R L G A L K S N I G H T Q A A A G V A G V I K M V Q A M R H G	GTTGAAGTCG L K S	AACATCGGGC N I G H	ACACGCAGGC T Q A	CGCGGCCGGT A A · G	GTCGCAGGTG V A G V	TCATCAAGAT I K M	GGTCCAGGCG V Q·A	ATGCGCCACG M R H G	28200
GACTGCT L L P	GACTGCTGCC GAAGACGCTG CACGTCGGA CCAGATCGAC TGGTCGGCGG GCACGGTGGA ACTCCTCACC GAGGCCGTCG ACTGGCCGGA 28300 L L P K T L H V D E P S D Q I D W S A G T V E L L T E A V D W P E	CACGTCGACG H V D E	AGCCCTCGGA P S D	CCAGATCGAC Q I D	TGGTCGGCGG WSAG	GCACGGTGGA T V E	ACTCCTCACC L L T	GAGGCCGTCG E A V D	ACTGGCCGGA W P E	28300
GAAGCAGGAC K Q D	GAAGCAGGAC GGCGGGCGGGC TGTCTCCTCC TTCGGCATCA GCGGGACGAA CGCGCACGTC GTCCTGGAGG AGGCCCCGGC GGTCGAGGAC 28400 K Q D G. G L R R A A V S S F G I S G T N A H V V L E E A P A V E D	GCCGCGCGC R A A	TGTCTCCTCC V S S	TTCGGCATCA F G I S	GCGGGACGAA G T N	CGCGCACGTC A H V	GTCCTGGAGG V L E E	AGGCCCGGC A P A	GGTCGAGGAC V E D	28400
TCCCGGCCG S P A V	TCCCCGGCCG TCGAGCCGCC GGCCGGTGCTGGTGC CGTGGCCGGT GTCCGGCGAAG ACTCCGGCCG CGCTGGACGC CCAGATCGGG CAGCTCGCCG 28500 S P A V E P P A G G G V V P W P V S A K T P A A L D A Q I G Q L A A	GGCCGGTGGC A G G	GGTGTGGTGC	CGTGGCCGGT W P V	GTCCGCGAAG S A K	ACTCCGGCCG T P A A	CGCTGGACGC L D A	CCAGATCGGG Q I G	CAGCTCGCCG Q L A A	28500

pikPKS Sequence

9	8	/	1	6	4	

C.	20	30	40	20	09	70	80	90	100	
1234567890	1234567890	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
CGTACGCGGA Y A D	CGGTCGTACG G R T	BamHI CGTACGCGGA CGGTCGTACG GACGTCGATC GCCGTACGGC GATGGAGCAC CGCGCGGTCG CGGTCGCCGA 28600 YAD GRID VDPAVAR ALVDSRIVDS RIAMEHRAVAVA VGD	CGGCGGTGGC A V · A	CGCCCGCGCC	GGTGGC CGCCCGCCC CTGGTCGACA GCCGTACGGC GATGGAGCAC CGCGCGGTCG CGGTCGGCG	GCCGTACGGC R T A	GATGGAGCAC M E H	CGCGCGGTCG R A V A	CGGTCGGCGA V G D	28600
CAGCCGGGAG S R E	GCACTGCGGG	CAGCCGGGAG GCACTGCGGG ACGCCCTGCG ATGCCGGAA GGACTGGTAC GCGGCACGTC CTCGGCGGGTGG CGTTCGTCTT CCCCGGCCAG 28700 S R E A L R D A L R M P E G L V R G T S S D V G R V A P V F P G Q	GATGCCGGAA M P E	GGACTGGTAC G L V R	GCGGCACGIC G I S	CTCGGACGTG S D V	GGCCGGGTGG G R V A	CGTTCGTCTT F V F	CCCCGGCCAG	28700
GGCACGCAGT G T Q W	GGCACGCAGT GGGCCGGCAT	Bemi GGCACGCAGT GGGCCGCAA CTCCTTGACA GCTCACCGGA GTTCGCTGCC TCGATGGCCG AATGCGAGAC CGCGCTCTCC CGCTACGTCG 28800 G T Q W A G M G A E L L D S S P E P A A S M A E C E T A L S R Y D	CTCCTTGACA L L D S	GCTCACCGGA S P E	GITCGCIGCC F A A	BSMI T STICGCIGCC ICGAIGGCCG AAIGCGAGAC CGCGCTCTCC F A A S M A E C E I A L S	Bemi AATGCGAGAC C E I	CGCGCTCTCC A L S	CGCTACGICG R Y V D	28800
ACTGGTCTCT W S L	r TGAAGCCGTC E A V	ACTGGICTCT IGAAGCCGIC GICCGACAGG AACCCGGGGC ACCCACGCGTCG ACGTCGICA GCCCGIGACC ITCGCTGTCA TGGICTCGCT 28900 W S L E A V V R Q E P G A P I L D R V D V V Q P V I F A V M V S L	AACCCGGCGC	ACCCACGCTC P T L	GACCGCGTCG D R V D	ACGTCGTCCA V V Q	GCCCGTGACC P V T	TTCGCTGTCA T F A V M Apall	TGGTCTCGCT V S L	28900

FIG. 31FFF

GGCGAAGGTC TGGCAGCACC ACGGCATCAC CCCCCAGGCC GTCGTCGGCC ACTCGCAGGG CGAGATCGCC GCCGCGTACG TCGCCGGTGC ACTCACCCTC 29000 A K V W Q H H G I T P Q A V V G H S Q G E I A A A Y V A G A L T L

99/1	64
------	----

-	GACGACGC CCCGCGTCGT CACCTGCGC AGCAAGTCCA TCGCCGCCCAGCGCGCCCAGGCGGCGAA TGATCTCCCT CGCCCTCGAC GAGGCGGCCG 29100 D D A A R V V T L R S K S I A A H L A G K G G M I S L A L D E A A V	TCCTGAAGCG ACTGAGCGAC TCTCCGTCGC CGCCGTCAAC GGCCCCACGCCGTCGT CTCCGGCGAC CCGACCCAGA TCGAGGAACT 29200 L K R L S D F D G L S V A A V N G P T A T V V S G D P T Q I E E L C C C C C C C C C C C C C C C C C C C	CGCCCGCACC TGCGAGGCCG ACGGCGCTCG TGCCCGGTCG ACTACGCCTC CCACAGCCGG CAGGTCGAGA TCATCGAGAA GGAGCTGGCC 29300 - A R T C E A D G V R A R I I P V D Y A S H S R Q V E I I E K E L A	⊕ GAGGICCTCG CCGGACGCT CCGCACGIGC CGITCTICTC CACCCTCGAA GGCACCTGGA TCACCGAGCC GGTGCTCGAC GGCACCTACT 29400 E V L A .G L A P Q A P H V P F F S I L E G I W I I E P V L D G I Y W	KpnI Acc65I ▼ ▼ GGTACCGCAA CCTGCGCCAT CGCGCCCGC CGTGGAGACC TTGGCGGTTG ACGGCTTCAC CCACTTCATC GAGGTCAGCG CCCACCCCGT 29500 GGTACCGCAA CCTGCGCCAT CGCGTGGGGCT TCGCCCCCGT 29500
	ISLA A	CTCCGGCGAC CCCS G D P	CAGGTCGAGA TC	TCACCGAGCC GG T B P V	CCACTTCATC GA H F I E
	AAGGGCGGCA 1 K G G M	CCACCGTCGT (T V V	CCACAGCCGG	GGCACCTGGA	ACGGCTTCAC G F T
	cctcccccc L A G	GGCCCCACCG G P T A	ACTACGCCTC Y A S	CACCCTCGAA T L E	TTGGCGGTTG L A V D
	TCGCCGCCA A A H	CGCCGTCAAC A V N	ATCCCGGTCG I P V D	CGITCITCIC F F S	CGTGGAGACC V E T
068/96#621	AGCAAGTCCA S K S I	TCTCCGTCGC S V A	TGCGCGGATC A R I PmlI	CCGCACGIGC P H V P	TCGCCCCGC A P A
1234567890	CACCCTGCGC T L R	TTCGACGGAC F D G L	ACGGCGTCCG G V R	CCCGCAGGCT P Q A	CGCGTGGGCT R V G . F
1234567890	CCCGCGTCGT R V V	ACTGAGCGAC L S D	TGCGAGGCCG C E A D	CCGGACTCGC .G L A	CCTGCGCCAT L R H
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	GACGACGCCG D D A A	TCCTGAAGCG L K R	CGCCCGCACC A R I	GAGGICCICG B V L A	KpnI Acc65I GGTACCGCAA Y R N

pikPKS Sequence

	29600	29700	29800	29900
1234567890	CCTCACCATG ACCCTCCCCG AGACCGTCAC CGGCCTCCGCC GCGAACAGGG AGGCCAGGAG CGTCTGGTCA CCTCACTCGC CGAAGCCTGG 29600 L T M T L P E T V T G L G T L R R E Q G G Q E R L V T S L A E A W	BCO47III GCCAACGGCC TCACCATCGA CTGGGCGCCC ATCTCCCCA CCACCCCCA CCTACGCCTT CCAGACCGAG CGCTTCTGGC 29700 A N G L T I D W A P I L P T A T G H H.P E L P T Y A F Q T E R F W L	PBtI SfcI ▼ ▼ TGCAGAGCTC CGCGCCCACC AGCGCCGCG ACGACTGGCG TTACCGCGTC GAGTGGAAGC CGCTGACGGCCAG GCGGACCTGT CCGGGGGTG 29800 TGCAGAGCTC CGCGCCCACC AGCGCCGCCG ACGACTGGCGTC GAGTGGAAGC CGCTGACGGCCAG GCGGACCTGT CCGGCGGTG 29800	GATCGTCGCC GȚCGGGAGCG AGCCAGAGGCGCGCGGG AGCGGAGGTC GACGTACTGG AAGCCGGGGC GGACGAC 29900 IVAVGSEPPEAELLGALKAAG AEVDVLEAGADCGGGG
90	CCTCACTCGC	CCAGACCGAG	GAGTGGAAGC CGCTGCGGCCAG GCGGACCTGT CE WKPLTASGQAD LS	AAGCCGGGGC
1234567890	S L A	Q T E		A G A
80	CGTCTGGTCA C	CCTACGCCTT	CTCCGGCCAG	GACGTACTGG
1234567890		Y A F	S G Q	D V L E
70	AGGCCAGGAG	GAGCTCCCCA	CGCTGACGGC	AGCGGAGGTC
1234567890	G Q E	E L P T	L T A	A E V
60	GCGAACAGGG	CCACCACCCC	GAGTGGAAGC	AGGCCGCGGG
1234567890	E Q G	H H · P	E W K P	A A G
50	CTCGGC ACCCTCCGCC G	CCGCAACCGG	TTACCGCGTC	GGCGCGCTGA
1234567890		A T G	Y R V	G A L K
40	CGGCCTCGGC	ATCCTCCCA	ACGACTGGCG	CGAGCTGCTG
1234567890	G L G	I L P T	D W R	E L L
30	AGACCGTCAC	CTGGGCGCCC	AGCGCCGCCG	AGCCAGAAGC
1234567890	T V T	W A P	S A A D	P E A
10 20 30 40 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	CTCACCATG ACCCTCCCCG AGACCGTCAC	TCACCATCGA T I D	PBtI T FCAGAGGTC CGCGCCACC AGCGCCGCG AGCGCCGC AGCGCCCGC AGCGCCCGC AGCGCCCGC AGCGCCCGC AGCGCCGCCG AGCGCCGCCG AGCGCCGCCG AGCGCCGC AGCGCCGCCG AGCGCCGCCG AGCGCCCGC AGCGCCCGCC	GŢCGGGAGCG V G S E
1234567890	CCTCACCATG L T M	GCCAACGGCC A N G L	SfcI TGCAGAGCTC Q S S	GATCGTCGCC I V A

pikPKS Sequence

FIG. 31HHH

CGTGAGGCCC TCGCCGCCCG GCTCACCGCA CTGACGACCG GCGACGGCTT CACCGGCGTG GTCTCGCTCC TCGACGACCT CGTGCCACAG GTCGCCTGGG 30000 REALAAR LIAAR LITAG DGF TGVVS LLDDL VPQ VAWV

	30100	30200	30300	30400
100 1234567890	TGCAGGCACT CGGCGACCGCC GGAATCAAGG CGCCCTGTG GTCCGTCACG TCTCCGTCGG ACGTCTCGAC ACCCCCGCCG ACCCCGACCG 30100	GGCCATGCTC TGGGGCTCG GCCGCGTCGT GGCCCGAAC GCTGGGCCGG CCTCGTCGAC CTCCCCGCCC AGCCCGATGC CGCCGTC 30200 A M L W G L G R V V A L E H P E R W A G L V D L P A Q P D A A A L	BBABI ▼ GCCCACCTCG TCACCGCACT CTCCGGCGCGAGG ACCAGATCGC CATCCGCACC ACCGGACTCC ACGCCCGCCGC CCTCGCCCGC GCACCCTCC 30300 A H L V T A L S G A T G E D Q I A I R T T G L H A R R L A R A P L H	ACGGACGICG GCCCACGCAGC CCCACGGCAC CGTCCTCATC ACCGGCGCA CCGGAGCCCT CGGCAGCCAC GCCGCACGCT GGATGGCCCA 30400 G R R P T R D W Q P H G T V L I T G G T G A L G S H A A R W M A H
90	ACCCCGCCG	AGCCCGATGC	CCTCGCCGC	GCCGCACGCT
1234567890	T P A D	P D A	L A R	A A R W
80	ACGTCTCGAC	CTCCCGCCC	ACGCCGGCG	CGGCAGCCAC
1234567890	R L D	L P A Q	A R R	G S H
70 1234567890	CCTGTG GTCCGTCACC CAGGGCGCGG TCTCCGTCGG ACGTCTCGAC ACCCCGCCG ACCCCGACCC	CACCCCGAAC GCTGGGCCGG CCTCGTCGAC CTCCCCGCCC AGCCCGATGC CGCCGCCC	ACCGGACTCC ACGCCCGCCG T G L H A R R	CCGGAGCCCT G A L
60 1234567890	CAGGGGGGG	GCTGGGCCGG WA'G	BBaBI r c catccgcacc I R T	C ACCGGCGCA C T G G T
50	GTCCGTCACC	сассссалас	BB * *GAGG ACCAGATCGC E D Q I A	CGTCCTCATC
1234567890	S V T	н р е в в		V L I
40	CGCCCTGTG P L W	CGCCCTTGAG	ACCGGCGAGG	CCCACGGCAC
1234567890		A L E	T G E D	H G T
30	GGAATCAAGG	eccecercer	CTCCGGCGCC	GACTGGCAGC
1234567890		R V V	S G A	D W Q P
	SCAGGCACT CGGCGACCC GGAATCAAGG CGCC	RECCATGETC TGGGGCCTCG GCCGCGTCGT	GCCCACCTCG TCACCGCACT CTCCGGCGCC ACCG A H L V T A L S G A T G	GRRP PTRDWQP
1234567890	TGCAGGCACT	GGCCATGCTC	GCCCACCTCG	ACGGACGTCG
	Q A L	A M L	A H L V	G R R

pikPKS Sequence

CCACGGAGCC GAACACCTCC TCCTCGTCAG CCGCAGCGGC GAACAAGCCC CCGAACCTCACC GCCGAACTCA CCGCATCGGG CGCCCGCGTC 30500 H G A E H L L L V S R S G E Q A P G A T Q L T A E L T A S G A R V

w ·

1 (<mark>)</mark> 2	/1	64
-----	------------------	----	----

10	20	30	40	50	09	70	80	96	100	
1234567890	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
ACCATCGCCG T I A A	ACCAICGCG CCIGCGACGI CGCCGACCC	CGCCGACCCC A D P	C CACGCCATGC H A M R	GCACCCTCCT T L L	GCACCCTCCT CGACGCCATC CCCCCCCTCAC CGCCGTCGTC CACACCGCCG	CCCGCCGAGA P A E T	CGCCCTCAC P L T	CGCCGTCGTC A V V	ACCATCGCCG CCTGCGACGC CACGCCATGC GCACCCTCCT CGACGCCATC CCCGCCGAGA CGCCCTCAC CGCCGTCGTC CACACCGCCG 30600 TIAAC CDVADPHAMRTLLDDAIPAETPLTAVVHTAG	30600
GCGCACCGGG A P G	RCACCGGG CGGCGATCCG CTGGACGTCA	CTGGACGTCA L D V T	E G P E	GGACATCGCC D I A	CCGGA GGACATCGC GGCATCCTGG GCGCGAAGAC GAGCGGCCCTCG ACGACCTGCT	GCGCGAAGAC A K T	GAGCGCCC S G A	GAGGICCICG E V L D	GCGCACCGGG CGGCGATCCG CTGGACGTCA CCGCCTCGCC CGCATCCTGG GCGCGAAGAC GAGCGGCGCC GAGGTCCTCG ACGACCTGCT 30700 A P G G D P L D V T G P E D I A R I L G A K T S G A E V L D D L L	30700
CCGCGGCACT R G T	CGCGGCACT CCGCTGGACG CCTTCGTCCT R G T P L D A F V L	CCTTCGTCCT F V L	CTACTCCTCG Y S S	AACGCCGGGG N A G V	AACGCCGGGG TCTGGGGCAG CGGCAGCCAG GGCGTCTACG CGGCGGCCAA N A G V W G S G S Q G V Y A A A N	CGGCAGCCAG	GGCGTCTACG G V Y A	CGGCGGCCAA A A N	CCGCGGCACT CCGCTGGACG CCTTCGTCCT CTACTCCTCG AACGCCGGGG TCTGGGGGCAG CGGCAGCCAG GGCGTCTACG CGGCGGCCAA CGCCCACCTC 30800 R G T P L D A F V L Y S S N A G V W G S G S Q G V Y A A A N A H L	30800
GACGCGCTCG D A L A	GACGCCTCG CCGCCCGG GGCG	CCGCGCCCGG R A R	GGCGAGACGG G E T A	ACGG CGACCTCGGT T A T S V	CGCCTGGGGC A W G	CTCTGGGCCG GCGACGGCAT L W A G D G M	GCGACGGCAT D G M	986000990 9 R G	MINI GACGCCCCGCCCCGCCCCGG GGCGAGACGG CGACCTCGGT CGCCTGGGGC CTCTGGGCCG GCGACGGCAT GGCCCGGGGC GCCGACGACG 30900 D A L A A R R R A R G E T A T S V A W G L W A G D G M G R G A D D A	1 30900
CGTACTGGCA	TACTGGCA GCGTCGGC ATCCGTCCGA	ATCCGTCCGA I R P M	TGAGCCCCGA S P D	CCGCGCCCTG R A L	MBCI BalI ▼ CCCGA CCGCGCCTG GACGAACTGG CCAAGGCCCT GAGCCACGAC GAGACCTTCG TCGCCGTGGC	MBCI Ball	GAGCCACGAC S H D	GAGACCTTCG E T F V	MBCI BALI ▼ CGTACTGGCA GCGTCCGA TGAGCCCCCGA CCGCGCCCTG GACGAACTGG CCAAGGCCCT GAGCCACGAC GAGACCTTCG TCGCCGTGGC 31000 Y W O R R G I R P M S P D R A L D E L A K A L S H D E T F V A V A	31000

pikPKS Sequence

GGCCGAACTG CCCCTCGACC 31500 A E L P L D R

103/164

	31100	31200	31300	31400
100	CGATGICGAC IGGGAGCGGI ICGCGCCCGC GITCACGGIG ICCCGICCA GCCIICTGCI CGACGGCGIC CCGGAGGCCC GGCAGGCGCI CGCCGCACCC 31100	GICGGIGCCC CGGCTCCCGG CGACGCCGCC GIGGCGCCGA CCGGGCAGTC GTCGGCGCTG GCCGCGATCA CCGCGGTCCC CGAGCCCGAG CGCCGGCGG 31200	CGCTCCTCAC CCTCGTCCGT ACCCACGCGG CGGCCGTACT CGGCCATTCC TCCCCCGAGC GGGTGGCCCC CGGCCGTGCC TTCACCGAGC TCGGCTTCGA 31300	CTCGCTGACG GCCGTGCAACCA GCTCTCCACG GTGGTCGGCA ACAGGCTCCC CGCCACCACG GTCTTCGACC ACCCGACGCC CGCCGCACTC 31400
1234567890	D V D W E R P A P I V S R P S L L L D G V P E A R Q A L A P P	V G A P A P G D A A V A P I G Q S S A L A A I I A L P E P E R R P A	L L T L V R T H A A A V L G H S S P D R V A P G R A F T E L G F D	S L T A V Q L R N Q L S T V V G N R L P A T T V F D H P T P A L
10 20 30 40 50 60 70 80 90 1034567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	ACGGTG TCCCGTCCA GCCTTCTGCT CGACGGCGTC CCGGAGGCCC GGCAGGCGCT CGCCGCA	CGAGCCCGAG E P E	TCCCCGACC GGGTGGCCCTTCACCGAGC TCGGCTTCS PDRVAPPGRAPCCTTC	ACCCGACGCC P T P
80 1234567890	CCGGAGGCCC	STCGCCGCTG GCCGCGATCA CCGCGCTCCC	CGGCCGTGCC G R A	GTCTTCGACC V F D H
70	CGACGGCGTC	GCCGCGATCA	GGGTGGCCCC	CGCCACCACG
1234567890	D G V	A A I T	V A · P	A T T
60	GCCTTCTGCT	GTCGGCGCTG	TCCCCCGACC	ACAGGCTCCC
1234567890	L L L	S A .L	S P D R	R L P
50	TCCCGTCCCA	CCGGGCAGTC	CGGCCATTCC	GTGGTCGGCA
1234567890	S R P S	G Q S	G H S	V V G N
40	GITCACGGIG	GTGGCGCCGA (CGGCCGTACT	GCTCTCCACG
1234567890	F I V	V A P T	A V L	L S T
30	TCGCGCCGC	CGACGCCGCC	ACCCACGCGG	TCCGCAACCA
1234567890	A P A	D A A	T H A A	R N Q
20 1234567890	GATGTCGAC TGGGAGCGGT TCGCGCCGC	GTCGGTGCCC CGGCTCCCGG CGACGCCGCC GTGG	CTCCTCAC CCTCGTCGT L L T L V R	GCCGTGCAGC A V Q L
1234567890	CGATGTCGAC	GTCGGTGCCC	CGCTCCTCAC	CTCGCTGACG
	D V D	V G A P	L L T	S L T

pikPKS Sequence

FIG. 31KKK

GCCGCGCACC TCCACGAGGC GTACCTCGCA CCGGCCGCAC GGACTGGGAG GGGCGGGTGC GCCGGGCCCT A A H L H E A Y L A P A E P A P T D W E G R V R R A L

1	04	/1	64
---	----	----	----

1034567890	10 20 30 40 50 60 70 80 90 100 100 100 100 100 100 100 100 100	30	40	50 1234567890	60 1234567890	70 1234567890	80 1234567890	90 1234567890	1234567890	
GGCTGCGGGA	GGCTGCGGGA CGCGGGGGTC CTCGACACCG TCCTGCGCCT CACCGGCATC GAGCCCGAGC CGGGTTCGGAC GGCGGCGCCG CCGACCCTGG 31600	CTCGACACCG	TCCTGCGCCT	CACCGGCATC	GAGCCCGAGC	CGGGTTCCGG	CGGTTCGGAC	ຄວວຄວຄວຄຄ	CCGACCCTGG	31600
I R U	гко в с с с в т	1 0 1	다 K	ы в	9 0 4 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	დ დ	က ဖ	ୟ ଷ ଓ	ტ <u>ი</u>	
TGCGGAGCCG A B P	TGCGGAGCCG GAGGCGTCGA TCGACGCCGAG GCCCTGATCC GGATGGCTCT CGGCCCCGT AACACCTGAC CCGACCGCGG TCCTGCCCCA 31700 A E P E A S I D D L D A B A L I R M A· L G P R N T	TCGACGACCT D D L	GGACGCCGAG D A B	G GCCTGAICC A L I R	GGATGCTCGGCCCCGG M A· L G P R	CGGCCCCGT G P R	T AACACCTGAC N T	cccacccccc	TCCTGCCCCA	31700
CGCGCCGCAC	CGCGCCGCAC CCCGCGCACCA CCCGCCCCCA CACGCCCACA ACCCCATCCA CGAGGGGAAG ACCACCACA GATGACGAGT TCCAACGAAC 31800 M T S S N E Q	CCGCGCACCA	CCCGCCCCCA	CACGCCCACA	ACCCCATCCA	саласавала	Accacaccca	GATGACGAGT M T S	r TCCAACGAAC S N E Q	31800
AGTTGGTGGA L V D	AGTTGGTGGA CGCTCTCTCA AGGAGAACGA AGAACTCCGG AAAGAGAGC GTCGCCGACCGTCGG CAGGAGCCCA TGGCGATCGT 31900 L V D A L R A S L K E N E E L R K E S R R A D R R Q E P M A I V Kpni Acc651	GCCTCTCTA A S L K	aggagaacga E N E	AGAACTCCGG E L R	GAACTCCGG AAAGAGACC GTCGCCGGGC CGACCGTCGG CAGGAGCCCA ELRKESRRRADRRQEPM ACC	GTCGCCGGC R R A	CGACCGTCGG D R R	CAGGAGCCCA Q E P M Acc	CA TGGCGATCGT M A I V KpnI Acc65I	31900
CGGCATGAGC G M S	CGGCATGAGC TGCCGGTTCG CGGGCGGAAT CCGGTCCCCC GAGGACCTCT GGGACGCCGT CGCCGGGGC AAGGACCTGG TCTCCGAGGT ACCGGAGGAG 32000 G M S C R F A G G I R S P E D L W D A V A A G K D L V S E V P E E	CGGGCGGAAT G G I	CCGGTCCCCC R S P	GAGGACCTCT E D L W	GGGACGCCGT D A V	CGCCGCGGGC A A G	AAGGACCTGG K D L V	TCTCCGAGGT S E V	ACCGGAGGAG P E E	32000

pikPKS Sequence

105/164

	32100	32200	32300	32400	32500
1234567890	CGCGGCTGGG ACATCGACTC CCTCTACGAC CCGCGAAGGG CACGACGTAC GTCCGCAACG CCGCGTTCCT CGACGACGCC GCCGGATTCG 32100	ACGCGGCTT CTTCGGGATC TCGCCGCGG AGGCCCTCGC CAGGAGCGGC AGCTCCTCGA AGCCTCCTGG GAGGTCTTCG AGCGGGGG 32200	CATCGACCCC GCGTCGGCACCGA CGTCGGCGTG TACGTGGGCT GTGGCTACCA GGACTACGCG CCGGACATCC GGGTCGCCCC CGAAGGCACC 32300	GECGGITACG TCGTCACCGG CAACTCCTCC GCCGTGGCCT CCGGGCGCAT CGCGTACTCC CTCGGCCTGG AGGGACCCGC CGTGACCGTG GACACGGCGT 32400	GCTCCTCTTC GCTCGTCGCC CTGCAAAGGG CCTGCGGAAC GGCGACTGCT CGACGGCACT CGTGGGCGGC GTGGCCGTCC TCGCGACGCC 32500
	R G W D I D S L Y D P V P G R K G T T Y V R N A A F L D D A A G F D	A A P F G I S P R E A L A M D P Q Q R Q L L E A S W E V F E R A G	I D P A S V R G T D V G V Y V G C G Y Q D Y A P D I R V A P E G T	G G Y V V T G N S S A V A S G R I A Y S L G L E G P A V T V D T A C	S S S L V A L H L A L K G L R N G D C S T A L V G G V A V L A T P
90	CGACGACGCC	GAGGTCTTCG	GGGTCGCCC	CGIGACCGIG	GTGGCCGTCC
1234567890	D D A	E V F E	V A P	V I V	V A V L
80	CCGCGTTCCT	AGCCTCCTGG	CCGGACATCC	AGGGACCCGC	cgraggcggc
1234567890	A F L	A S W	P D I R	G P A	V G G
70	GTCCGCAACG	AGCTCCTCGA	GGACTACGCG	CTCGGCCTGG	CGACGGCACT
1234567890	V R N A	L L E	D Y A	L G L E	T A L
60	CACGACGTAC	CAGCAGCGGC	GTGGCTACCA	CGCGTACTCC	GGCGACTGCT
1234567890	T T Y	Q Q R Q	G Y Q	A Y S	G D C S
50	GGCGCAAGGG	CATGGACCCG	TACGTGGGCT	CCGGGCGCAT	CCTGCGGAAC
1234567890	R K G	M D P	Y V G C	G R I	L R N
40	b d A d	AGGCCCTCGC	CGTCGGCGTG	GCCGTGGCCT	CCCTGAAGGG
1234567890		A L A	V G V	A V A S	L K G
30	ccrcracgac	TCGCCGCGCG	GCGGCACCGA	CAACTCCTCC	CTGCACCTCG
1234567890	L Y D	S P R E	G T D	N S S	L H L A
10 20 30 40 50 60 70 80 90 100 100 100 100 1234567890 1	ACATCGACTC	CTTCGGGATC	GCGTCGGTCC	TCGTCACCGG	GCTCGTCGCC
	I D S	F G I	A S V R	V T G	L V A
1234567890	CGCGGCTGGG	ACGCGGCCTT	CATCGACCCC	GGCGGTTACG	GCTCCTCTTC
	R G W D	A A F	I D P	G G Y V	S S S

pikPKS Sequence

106/164

o F	20	30	40	50	09	70	80	06	100	
1234567890	12345678	1234567890	1234567890	1234567890	1234567890	67890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	
GGCGCGTTC	GGGCGCTTC ATCGAGTTCA GCAGCCAGCC GCCGACGGCC GGACCAAGGG CTTCGCCTCG GCGGCGGACG GCCTCGCCTG GGGCGAGGC 32600 G A F I E F S S Q Q A M A A D G R T K G F A S A A D G L A W G E G	S Q Q	GGCCATGGCC A M A	GCCGACGGCC A D G R	GGACCAAGGG T K G	CTTCGCCTCG F A S	GCGGCGGACG A A D G	GCCTCGCCTG L A W	GGCCGAGGCC	32600
GTCGCCGTAC V A V L	GTCGCCGTAC TCCTCCTCGA ACGCCTCTCC GACGCGGCGGCCA CCGGGTCCTG GCCGTCGTGC GCGGCAGCGC CATCAACCAG GACGGCGA 32700 V A V L L E R L S D A R R K G H R V ·L A V V R G S A I N Q D G A S	ACGGCTCTCC R L S	GACGCGCGC D A R R	GCAAGGGCCA K G H	CCGGGTCCTG R V ·L	GCCGTCGTGC A V V R	GCGGCAGCGC GSA	CATCAACCAG I N Q	gacggcga D g a s	32700
GCAACGGCCT N G L	GCAACGGCCT CACGGCCCT CCCAGCAGCA CCTGATCCGC CAGGCCCTGG CCGACGCGCG GCTCACGTCG AGCGACGTGG ACGTCGTGGA 32800 N G L T A P H G P S Q Q H L I R Q A L A D A R L T S S D V D V V E	CACGGGCCT H G P S	CCCAGCAGCA Q Q H	CCTGATCCGC L I R	CAGGCCCTGG Q A L A	CCGACGCGCG D A R	GCTCACGTCG L T S	AGCGACGTGG S D V D	ACGTCGTGGA V V E	32800
GGGCCACGGC G H G	ABCI GGGCCACGGC ACGGGGACCC GTCTCGGCGA CCCCAGGCGC TGCTCGCCAC GTACGGGCAG GGGCGCCC CGGGGCAGCC GCTGCGGCTG 32900 G H G T G T R L G D P I E A Q A L L A T Y G Q G R A P G Q P L R L	Greregega L G D	CCCGATCGAG P I E	GCGCAGGCGC A Q A L	TGCTCGCCAC L A T	GTACGGGCAG Y G Q	AscI GGGCGCCC GRAP	CGGGGCAGCC G Q P	GCTGCGGCTG L R L	32900
GGGACGCTGA G T L K	gggacgctga agtcgaacat cgggcacacacg caggcgctt cgggtgtcgc cggtgtcatc aagatggtgc aggcgctgcg ccacggggtg ctgccgaaga 33000 g t l k s n i g h t q a a s g v a g v i k m v q a l r h g v l p k t	CGGGCACACG G H T	CAGGCCGCTT Q A A S	CGGGTGTCGC G V A	CGGTGTCATC G V I	AAGATGGTGC K M V Q	AGGCGCTGCG A L R	CCACGGGGTG H G V	CTGCCGAAGA L P K T	33000

pikPKS Sequence

1234567890	20 1234567890	30 1234567890	10 20 30 40 50 60 70 80 100 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	50 1234567890	60 1234567890	70 1234567890	80 1234567890	90 1234567890	1234567890	
Pml I CCCTGCACGT G	1.I GGACGAGCCG	ACGGACCAGG	TCGACTGGTC	DOLLOSOODS	GTCGAGCTGC	TCACCGAGGC	CGTGGACTGG	CGGGAGCGGC	PmlI CCTICAACT GAACAACCA ACGGACCAGG TCGACTGGTTCG GTCGAGCTGC TCACCGAGGC CGTGGACTGG CCGGAGCGGC CGGGCT 33100	33100
V H I	D B P	T D Q V	D W S	A G S AACGCGCACG	V E L L TCGTCCTGGA	T E A GGAGGCCCG	V D W	P B K F AGTCCCTGC	L H V D E P T D Q V D W S A G S V E L L T E A V D W F E L L T E A O D W F E K F G K L CCRCCGGGG G G G G G G G G G G G G G G G	33200
ж 4	G V S . A	RRAGO SAFGV	T D D	V A H V	G T N A H V V L E E A P A V E E S P A V E P	E A P	A V E E GGCAGCTCGC	S P A	V E P GAAGACCGCA	33300
CCGGCGGTG P.A.G.G Bal	TG GCGGCGTGGT G G V V BamHI	P W P	V S A K	T S A	A L D	A Q I G	4 1 0	A Y A	PAGGGGGG GCGGCGGGGGGGGGGGGGGGGGGGGGGGGGG	,
CGGACGTGGA D V D	rccccccrc P A V	HANCGTGGA TCCGGCGGTG GCCGCCCGCG CCCTD V D P A V A A R A L	CCCTGGTCGA L V D	CAGCCGTACG S R T	Y D S'R T A M E H R A V A V G D S R E A L R	ACCGCGCGGT R A V	CGCGGTCGGC A V G	GACAGCCGGG D S R E	CGGACGTGGA TCCGGCGCGCCCGCG CCCTGGTCGA CAGCCGTACG GCGATGGAGC ACGCGGGT CGCGGTCGGC GACAGCCGGG AGGCACTGCG 33400 D v d p a v a a r a l v d s r t a m e h r a v a v g d s r e a l r	33400

pikPKS Sequence

FIG. 31000

GGACGCCTG CGGATGCCGG AAGGACTGGT ACGGGGCACG GTCACCGATC CGGGCCGGGT GGCGTTCGTC TTCCCCGGCC AGGGCACGCA GTGGCCGGC 33500 D A L R M P E G L V R G T V T D P G R V A F V F P G Q G T Q W A G

108/164	
---------	--

10	20 1234567890	10 20 30 40 50 60 70 80 90 1000 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	40 1234567890	50 1234567890	60 1234567890	70 1234567890	80 1234567890	90 1234567890	1234567890	
ည္ အ	ATGGGCGCG AACTCCTCGA CAGCTCACC	CAGCTCACCC	ECORI APOI GAATTCGCCG EFAA	I CGCCG CCGCCATGGC A A A M A	BemI CGAATGCGAG A(E C E T	Bemi TGAATGCGAG ACCGCACTCT CCCCGTACGT CGACTGGTCT CTCGAAGCCG E C E T A L S P Y V D W S L E A V	CCCGTACGT P Y V	CGACTGGTCT D W S	ECORI APOI T ATGGCGCCG AACTCCTCGA CAGCTCACCC GAATTCGCCG CGAATGCGAG ACCGCACTCT CCCGTACGT CGACTGGTCT CTCGAAGCCG 33600	3600
ජී ය	GGCTCCCAGC A P S.	GCACCGACAC A P T L	TCGACCGCGT D R V	CGACGTCGTC D V V	CAGCCCGTCA Q P V T	CCTTCGCCGT F A V	CATGGTCTCC M V S	CTCGCCAAGG L A K V	TCGTCCGACA GGCTCCCAGC GCACCGACAC TCGACCGCTC CAGCCCGTCA CCTTCGCCGT CATGGTCTCC CTCGCCAAGG TCTGGCAGCA 33700	3700
CACGCCATC H G I	ACCCCCGAGG T P E A	ACCCCCGAGG CCGTCATCGG T P E A V I G	CCACTCCCAG H S Q	GGCGAGATCG G E I A	GGCGAGATCG CCGCCGCGTA G E I A A A Y	CGTCGCCGGT GCCCTCACCC TCGACGACGC	GCCCTCACCC A L T L	TCGACGACGC D D A	CCACGGCATC ACCCCCGAGG CCACTCCCAG GGCGAGATCG CCGCCGCGTA CGTCGCCGGT GCCCTCACCC TCGACGACGC CGCTCGTGTC 33800 H G I T P E A V I G H S Q G E I A A A Y V A G A L T L D D A A R V	3800
ပ္ပ် 🛪	GTGACCCTCC GCAGCAAGTC (CATCGCCGCC I A A	CACCTCGCCG H L A G	GCAAGGGCGG K G G	CATGATCTCC M I S	CATGATCTCC CTCGCCCTCA G M I S L A L S	GCGAGGAAGC E E A	CACCCGGCAG T R Q	GTGACCTICC GCAGCAAGTC CATCGCCGC CACGCGGCGGCGG CATGATCTCC CTCGCCCTCA GCGAGGAAGC CACCCGGCAG CGCATCGAGA 33900 V T L R S K S I A A H L A G K G G M I S L A L S E E A T R Q R I E N	13900

pikPKS Sequence

FIG. 31PPP

ACCTCCACGG ACTGTCGATC GCCGCCGTCA ACGGGCCTAC CGCCACCGTG GTTTCGGGCG ACCCCACCCA GATCCAAGAA CTTGCTCAGG CGTGTGAGGC 34000 L H G L S I A A V N G P T A T V V S G D P T Q I Q E L A Q A C E A

1234567890	20	30	40 1234567890	50 1234567890	60 1234567890	70 1234567890	80 1234567890	10 234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	100 234567890
CGACGGCATC D G I	CGCGCACGGA R A R I	TCATCCCGT I P V	CGACTACGCC D Y A	TCCCACAGCG S H S A	CCCACGTCGA H V E	GACCATCGAG T I E	AACGAACTCG N E L A	CCGACGTCCT GO	CGACGCCATC CGCGCACGGA TCATCCCCGT CGACTACGCC TCCCACGTCGA GACCATCGAG AACGAACTCG CCGACGTCT GGCGGGTTG 34100 D G I R A R I I P V D Y A S H S A H V E T I E N E L A D V L A G L
TCCCCCCAGA CAS P Q T	CACCCCAGGT	ACCCCAGGT CCCCTTCTTC P Q V P F F	TCCACCCTCG	AAGGCACCTG G T W	GATCACCGAA I T E	CCCGCCCTCG	ACGGCGGCTA G G Y	KpnI Acc65I	Kpn1 Acc651 TCCCCCCAGA CACCCCAGGT CCCCTTCTTC TCCACCCTCG AAGGCACCTG GATCACGGAA CCCGCCCTCG ACGGCGGCTA CTGGTACCGC AACCTCCGCC 34200 S P Q T P Q V P F F S T L E G T W I T E P A L D G G Y W Y R N L R H
ATCGTGTGGG	crrcccccs F A P	GCCGTCGAGA A V E T	CCCTCGCCAC L A T	CGACGAAGGC D E G	TTCACCCACT F T H F	TCATCGAGGT I E V	CAGCGCCCAC S A H	CCCGTCTCA C	ATCGIGIGGG CTTCGCCCCG GCCGTCGAGC CGACGAAGGC TTCACCCACT TCATCGAGGT CAGCGCCCAC CCCGTCCTCA CCATGACCCT 34300 R V G F A P A V E T L A T D E G F T H F I E V S A H P V L T M T L
CCCCGACAAG P D K	Gricaccesce T	MSCI Bali TGCCACCCT	CCGACGCGAG R R E	GACGGCGGAC D G G Q	AGCACCGCCT H R L	ACGCGAG GACGGCGAC AGCACCGCT CACCACCTCC CTTGCCGAGG CCTGGGCCAA R E D G G Q H R L T T S L A E A W A N	CTTGCCGAGG L A E A	CCTGGGCCAA O	MBGI Bali ▼ CCCCGACAAG GTCACCGGCC TGGCCACCCGAG GACGGCGGAC AGCACCGCT CACCACCTCC CTTGCCGAGG CCTGGGCCAA CGGCCTCGCC 34400 P D K V T G L A T L R R E D G G Q H R L T T S L A E A W A N G L A
CTCGACTGGG	F CTCCCTCCT	GCCCGCCACG P A T	GGCGCCCTCA G A L S	GCCCGCCGT P A V	CCCCGACCTC P D L	CCGACGTACG P T Y A	CCTTCCAGCA F Q H	CCGCTCGTAC TO R S Y W	CTCGACTGGG CCTCCCTCCT GCCCGCCCTCA GCCCCGCCGT CCCCGACCTACG CCTTCCAGCA CCGCTCGTAC TGGATCAGCC 34500 L D W A S L L P A T G A L S P A V P D L P T Y A F Q H R S Y W I S P

pikPKS Sequence

	34600	34700	34800 ACP 6	34900 domain
100 1234567890	CCGCGGGTCC CGGCGAGGC CCGCGCACA CCGCTTCCGG GCGCGAGGCC GTCGGGGCCCG GGTGCCGAGG ACCTCGACGA 34600 A G P G E A P A H T A S G R E A V A E T G L A W G P G A E D L D E	GGAGGCCGG CGCAGCGCCG TACTCGCGAT GGTGATGCGG CAGGCGCCT CCGTCCG GTGCGACTCG CCCGAAGAGG TCCCCGTCGA CCGCCCGCTG 34700 E G R R S A V L A M V M R Q A A S V L·R C D S P E E V P V D R P L	CGGGAGATCG GCTTCGACTCGCC GTCGACTTCC GCAACCGCGT CAACCGGCTG ACCGGTCTCC AGCTGCCGCC CACCGTCGTG TTCCAGCACC 34800 R E I G P D S L T A V D F R N R V N R L T G L Q L P P T V V F Q H P ACP 6	CGACCCCGT CGCGCTCGCC GAGCGCTCA GCGACGAGCGG AACTGGGCCG TCGCCGAGCC GTCGGATCAC GAGCAGCGG AGGAGGAGAA 34900 TPVALLAERISDELAERNWAVAEPSDHESQAECG TCGCCGAGCC GTCGGATCAC GAGCAGCCGG AGGAGGAGAA 34900 TE domain
10 20 30 40 50 60 70 80 90 103	GCGCGAGGCC GTCGCCGAGA CGGGGCTCGC GTGCCCGAGG ACCTCGAC	TCCCCGTCGA	CACCGTCGTG	GAGCAGGCGG
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	R E A V A E T G L A W G P G A E D L D	P V D	T V V	E Q A E
80	GTGGGGCCCG	CAGGCGCCT CCGTGCTCG GTGCGACTCG CCCGAAGAGG TCCCCGTCGA	AGCTGCCGCC	GTCGGATCAC
1234567890	W G P		L P P	S D H
70	CGGGGCTCGC	GTGCGACTCG	CAACCGGCTG ACCGGTCTCC AGCTGCCGCC	AACTGGGCCG TCGCCGAGCC ON WAN AN AE P
1234567890	G L A	C D S	N R L T G L Q L P P	
60	GTCGCCGAGA	CCGTGCTCCG	CAACCGGCTG	AACTGGGCCG
1234567890	V A E T	V L·R	N R L	N W A V
50	GCGCGAGGCC	CAGGGGCCT	GCAACCGCGT	GGCCGAGCGG
1234567890	R E A	Q A A S	N R V	A B R
40	CCGCTTCCGG	GGTGATGCGG	GTCGACTTCC G	GCGACGAGCT
1234567890	A S G	V M R	V D F R	D E L
30	CGGCGAGGC CCCGCGCACA	TACTCGCGAT	GCTGACCGCC	CGCGCTCGCC GAGCGCATCA
1234567890		L A M	L T A	A L A E R I S
20	CGGCGAGGCG	GAGGGCGG CGCAGCGCG TACTCGCGAT GGTGA	CGGGAGATCG GCTTCGACTC GCTGACCGCC GTCGA	CGCGCTCGCC
1234567890	G E A	E G R R S A V L A M V M	R E I G F D S L T A V D	A L A
10	CCGCGGGTCC	GGAGGGCCGG	CGGGAGATCG	CGACGCCGT
1234567890	A G P	E G R	R E I G	T P V

pikPKS Sequence

FIG. 31RRR

GGCCGCCGT CCGCCGCGGGGG CCCGCTCCGG GGCCGCCGGCG CCGGGATGTT CCGCGCCCTG TTCCGGCAGG CCGTGAGGA CGACCGGTAC 35000

111/164

0	20	30	40	50	9	70	80	96	100	
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	
GGCGAGTTCC G E F L	GGCGAGTTCC TCGACGTCCT CGCCGAAGCC TCCG	CGCCGAAGCC A E A	TCCGCGTTCC S A F R	GCCCGCAGIT P Q F	CGCCTCGCCC A S P	GGCGAGTTCC TCGACGTCC TCCGCGTTCC GCCCGCAGTT CGCCTCGCCC GAGGCCTGCT CGGAGCGGCT CGACCCGGTG CTGCTCGCCG 35100	CGGAGCGGCT E R L	CGACCCGGTG D P V	CTGCTCGCCG L L A G	35100
GCGGTCCGAC G P T	GGACCGGGCG D R A	GAAGGCCGTG E G R A	CCGTTCTCGT V L V	CGGCTGCACC G C T	GGCACCGCGG	EBPI BPU11021 F GCGTCCGAC GGACCGGGC GAAGGCCGTG CCGTTCTCGT CGGCTGCACC GGCACCGCG CGAACGGCGG CCCGCACGAG TTCCTGCGGC TCAGCACCTC 35200 G P T D R A E G R A V L V G C T G T A A N G G P H E F L R L S T S	CCCGCACGAG P H B	TTCCTGCGGC F L R L	Espi Bpull02I TCAGCACCTC S T S	35200
CTTCCAGGAG P Q E	GAGCGGGACT TCCTCGCCG	TCCTCGCCGT L A V	ACCTCTCCC P L P	GGCTACGGCA	CGGGTACGGG G T G	CTTCCAGGAG GAGCGGGART TCCTCGCGT ACCTCTCCC GGCTACGGC CACCGGCACG GCCCTCCTCC CGGCCGATCT CGACACCGCG 35300 F Q E E R D F L A V P L P G Y G T G T G T G T A L L P A D L D T A	GCCCTCCTCC A L L P	CGGCCGATCT A D L	CGACACCGCG D T A	35300
CTCGACGCCC L D A Q	CTCGACGCCC AGGCCCGGGC	GATCCTCCGG I L R	GCCGCCGGGG A A G D	ACGCCCGGT A P V	CGICCIGCIC V L L	CTCGACGCCC AGGCCCGGGC GCCGCCGGGG ACGCCCCGGT CGTCCTGCTC GGGCACTCCG GCGGCGCCCT GCTCGCGCAC GAGCTGGCCT 35400 L D A Q A R A I L R A A G D A P V L L G H S G G A L L A H E L A F	GCGGCGCCT	GCTCGCGCAC L A H	: GAGCTGGCCT E L A F	35400
TCCGCCTGGA (GCGGGCGCAC R A H	Asci GGCGCCCCC G A P P	CGGCCGGGAT A G I	CGTCCTGGTC V L V	GACCCCTATC D P Y P	Asci TCCGCCTGGA GCGGGCGC GGCCCGGCAT CGTCCTGGTC GACCCCTATC CGCCGGGCCA TCAGGAGCCC ATCGAGGTGT GGAGCAGGCA 35500 R L E R A H G A P P A G I V L V D P Y P P G H Q E P I E V W S R Q	TCAGGAGCCC Q E P	SCCC ATCGAGGTGT P I E V W	GGAGCAGGCA S R Q	35500

pikPKS Sequence

ъ С С

O

လ

A L

P S 4

田田田

R F S

Œ

S

א ט. ט

H A G

n u

10	20	30	40	50	.09	70	80	06	100	
1234567890	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
อหออออฉรอ	MBCI Bali	CGGGCGAGCT	GGAGCCGATG	TCCGATGCGC	MBCI Bali ▼ GGCTGCTGGC CA	cI ll CATGGGCCGG M G R	TACGCGCGGT	TCCTCGCCGG	SCCGCGCGGCG	35600
GGCCGCAGCA C	LGEGGGGGCCCGT GCTTCTGGTC CGTGCCTCCG AACCGCTGGG CGACTGGCAG GAGGAGCGGG GCGACTGGCG TGCCCACTGG GACCTTCCGC 35700 GGCCGCAGCA GCGCGCCCGT GCTTCTGGTC CGTGCCTCCG AACCGCTGGG CGACTGGCAG GAGGAGCGGG GCGACTGGCG TGCCCACTGG GACCTTCCGC 35700 GRSSAPVLLVRASBPLGDWQ BERG DWKAAHWDLPH	G L F A G E L C GCGCCCGT GCTTCTGGTC A P V L L V	CGTGCCTCCG	AACCGCTGGG	CGACTGGCAG D W Q	GAGGAGCGGG G	GCGACTGGCG	TGCCCACTGG	GACCTTCCGC D L P H	35700
ACACCGTCGC T V A	ACACCGTCGC GGACGTGCACT TCACGATGAT GCGGGACCAC GCGCCGGCCG TCGCCGAGGC CGTCCTCTC TGGCTCGACG CCATCGAGGG 35800 T V A D V P G D H F T M M R D H A P A V A E A V L S W L D A I E G	GGCGACCACT G D H F	TCACGATGAT T · M M	GCGGGACCAC R D H	GCGCCGGCCG 1	TCGCCGAGGC A E A	CGTCCTCTCC V L S	TGGCTCGACG W L D A	CCATCGAGGG I E G	35800
CATCGAGGG	BamHI	* GACCGACAGA	CCTCTGAACG	TGGACAGGG	BamHI A ACTGTGGATC	ii ceacacticc	ACCCCGCGCC	GAACAGCGCG	GTGCGGCTGG	35900
ъ М	> ↑ ⊻ ♥	T C	P L N V	დ დ	H M	я 4 4	а «	N N	VRLV	TE2

pikPKS Sequence

FIG. 31TTT

TOTGCOTGCC GCACGCCGGC GGCTCCGCCA GCTACTTCTT CCGCTTCTCG GAGGAGCTGC ACCCCTCCGT CGAGGCCCTG TCGGTGCAGT ATCCGGGCCG 36000

CLRA

G P F

E E

33 23 23

Ω,

D P K A

G U R

ALA

CPVM

113/164

0.5	30	30	40	50	9	70	80	06	100
1234567890		1234567890 1234567890		1234567890 1234567890 1234567890 1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CCAGGACCGG	CGTGCCGAGC	CGTGTCTGGA	GAGCGTCGAG	GAGCICGCCG	AGCATGTGGT	CGCGGCCACC	свыссствет	CCAGGACCGG CGTGCCGAGC CGTGTCTGGA GAGCGTCGAG GAGCTCGCCG AGCATGTGGT CGCGGCCACC GAACCCTGGT GGCAGGAGGG CCGGCTGGCC 36100	GCTGGCC 36100
О В	ж я 9	C I	В С	я к я	ΛΛΗ	A A T	E P W W	о В	r A
TTCTTCGGGC	ACAGCCTCGG	CGCCTCCGTC	GCCTTCGAGA	CGGCCCGCAT	CCTGGAACAG	CGGCACGGGG	TACGGCCCGA	TICTICAGAC ACAGCCICGA CACCICCGIC GCCTICGAGA CAGCCCGCAI CCTGGAACAG CGGCACGAGG IACGGCCCGA GGGCCTGIAC GICTCCGGIC 36200	rccecrc 36200
H 0 4	STG	N S	A F E	A R H	л О	в н с	स स	D Y J	м С
₽	*						Espi Bpul	Espi Bpull021	
2222525255 ▲	GTCGCTGGCG	CCGGACCGGC	TCGTCCACCA	GCTGGACGAC	CGGGCGTTCC	TGGCCGAGAT	► CCGGCGGCTC	GOCCCCCC GICCTIGGC CCGGACCGC ICGICCACCA GCIGGACGAC CGGGCGIICC IGGCCGAGAI CCGGCGGCIC AGCGGCACCG ACGAGCGGII 36300	AGCGGTT 36300
R A P	ა 1	PDRU	О н л	L D	R A F L	A E I	ж 1	S G T D	E4 CX
CCTCCAGGAC	GACGAGCTGC	TGCGGCTGGT	BCTGCCCGCG	CTGCGCAGCG	ACTACAAGGC	GCGGAGACG	TACCTGCACC	CCTCCAGGAC GACGAGCTGC TGCGGCTGGT GCTGCCCGCG CTGCGAGCG ACTACAAGGC GGCGGAGACG TACCTGCACC GGCCGTCCGC CAAGCTCACC 36400	AGCTCACC 36400
L Q D	DELL	R L V	L P	L R S D	Y K A	A E	YLHR	P S A K	타
TGCCCGGTGA	TGGCCCTGGC	CGGCGACCGT	GACCCGAAGG	CGCCGCTGAA	CGAGGTGGCC	GAGTGGCGTC	GGCACACCAG	reccesta tescettese escesacet sacecsaags escesetsaa csassissee sasteseete sseacaeeas essecette tseeteess 36500	cercede 36500

pikPKS Sequence

FIG. 31000

36778

				,	
		BGIII		CGTCGTGCAG CCCCCGACCA GCCTTATCGA AGGAGCGGCG AGAAGATGGC AGAAJCCCACG GTGACCGACG ACCTGACGG GGCCCTCACG CAGCCCCCGC 36700	
100	90	900	~	អ្ន	
	5678	ນວອມ	ez er	ij	
	234	CGA	Ω	A GC	
90	0 13	, Ď		. 0	
6	789	AscI (Д	SE SE	
	1456	₽	ধ ত	9	
	123	g		ğ	
80	890	ACC	e E	99	
	567	GTC	.	DE SEC	
	234	CTC	1 >	ä	•
5	100	. K		គ្គ	
	6789	CAC	т н о	CGA	
	345	GAC	Ħ	GAG	
	12	8	Α	텀	
9	1890	VTCT	w	80	æ
	1567	340	D G	ည်	Ω
	123,	AAC	A	AGA	Z
20	90	၂ ၁၉	z	ာမ္မ	œ
	678	Bglll P SATCTG	U	PATG	3
	2345	P B B	H	3AA(R
l	;; 0	Ö 4	মে	₹	×
64	789	299	Ħ	ပ် ဗိ	æ
	456	AGT	æ .œ	P GC	æ
	123	₩ CC	α.	A G	ტ
30	890	Acc	Ω	49 5	四
	567	វិ	r Z	TAT	H
	1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890	TACC		GCCI	H
20	06	ŢĊ,	H +	ð	တ
	678	.ac.	[te ₄	GAC	E G
	2345	999	# *	ggg	ф
	0 1;	8 6	Q	Ŭ ტ	
P	789	85 85	ರ	3	ø
	1456	ACT	დ ჯ	CGT	0 0 0
	123	S	₩	55	>

pikPKS Sequence

FIG. 31WV

TREGCCACIAC CATCCACACA GIGACCAACA GIGAACTCAG CACCCACCIC CIAGAGACCC GCGGCAICCA CIGGAICC

10 20 30 40 50 60 70 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	70 80 90 3456789012345678901234567890	
BamHI ▼ GGATCCGGCGCTTCCACCCCGCCGCGCGCGCGCTGCCTGC	GGCGGCTCCGCTACTTCTTCCGCT CCGCCGAGGCGGTCGATGAAGAAGGCGA G S A S Y F F R F	06
TCTCGGAGGAGCTGCACCCCTCCGTCGAGGCCCTGTCGGTGCAGTATCCGGGCCGCCAGGACCCGGCGTGCCGAGCCGTGTCTGGAGAGCG AGAGCCTCCTCGACGTGGGGAGGCAGCTCCGGGACAGCCACGTCATAGGCCCGGCGGCGGTCCTGGCCGCACGGCTCGGCACAGACCTCTCGC SEELHPSVE ALSONES OF SON EN LON ON PGRONE ON RABPC SON	CGGCGTGCCGAGGTGTCTGGAGAGCG SCCGCACGGCACAGACCTCTCGC R R A E P C L E S V	180
N8PHI ▼ TCGAGGAGCTCGCCGAGCATGTCGCGGCCACCGAACCCTGGTGGCAGGGCCGGCTGGCCTTCTTCGGGCACAGCCTCGGCGCCT AGCTCCTCGAGCGGCGTCGTACACCGGCGGCGGGGGCCGGCC	SCCTTCTTCGGGCACAGCCTCGGCGCCT CGGAAGAAGCCGTGTCGGAGCCGCGGA A F F G H S L G A S	270
ABOI CCGTCGCCTTCGAGACGGCCCGCATCCTGGAACAGCGGCACGGGCTACGGCCCTGTACGTCTCCGGTCGGCGCCCCCGTCGC GGCAGCGGAAGCTCTGCCGGGCCTTGTCGCCGTGCCCCATGCCGGGCCTGTACGTCTCCGGTCGGCCGCGCGCG	AscI TACGTCTCCGGTCGCCCCCCTCGC ATGCAGGCCAGCCGCGCGCGCAGCG	360

116/164

10 20	30	40	50	09	70	08	8	
123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890	567890123456	7890123456	7890123456	789012345	6789012345	678901234	567890	
RerII	Pvuli PflMI			ES GD	Espi Bpull02I			
TGGCGCCGGACCGGCTCCACCAGCTGGACGGCGTTCCTGGCCGAGGCGCGCGC	AGCTGGACGACCC FTCGACCTGCTGG	GGGCGTTCCT CCCGCAAGGA A F L	GGCCGAGATC CCGGCTCTAC A E I	CGGCGGCTC	GACCGGGCGTTCCTGGCCGAGGCGCTCAGCGGCACCGACGAGGGGTTCCTCC SCTGGCCCGCAAGGACCGGCTTAGGCCGCCGCGAGGTCGCCGTGGCTGCTCGCCAAGGAGG D R A F L A E I R R L S G T D E R F L Q	ACGAGCGGT TGCTCGCCA E R F	TCCTCC AGGAGG L Q	450
		FspI						
AGGACGACGAGCTGCTGCTGCTGCCCGCGCTGCGCAGCGACTACAAGGCGGCGGAGACGTACCTGCACCGGCCGTCCGCCAAGC TCCTGCTGCTCGACGACGACGACGGCGCGCGCGCGTCGCTGATGTTCCGCCGCCGCTTTGGACGTGGACGTGGCCGGCC	FIGCTGCCCGCGCTGCGC PACGACGGGCGCGACGCG 7 L P A L R	TGCGCAGCGA ACGCGTCGCT R S D	CTACAAGGC GATGTTCCG(Y K A	GCGGAGACG CCGCCTCTGC A E T	AGCGACTACAAGGCGGCGGAGACGTACCTGCACCGGCCGTCCGCCAA(STCGCTGATGTTCCGCCGCCTCTGCATGGACGTGGCCGGCAGGCGGTT(S D Y K A A E T Y L H R P S A K	GGCCGTCCG CCGGCAGGC PSA	CCAAGC GGTTCG K L	540
					•		,	,
TCACCTGCCCGGTGATGGCCCTGGCCGGCGACCGTGAAGGCGCCGCTGAACGAGGTGGCCGAGTGGCGTCGGCACACCAGCGGGC AGTGGACGGGCCACTACCGGGACCGGCCGCTGGCTTCCGCGGCGACTTGCTCCACCGGCTCACCGCAGTGGCCGTGTGGTCGCCC T C P V M A L A G D R D P K A P L N E V A E W R R H T S G P	SCCGGCGACCGTG SGCCGCTGGCAC'S A G D R D	CARCCGAAGGCGCCGCTGAACGAGTGGCCGAGTGC CTGGGCTTCCGCGGCGACTTGCTCCACCGGCTCACC D P K A P L N E V A E W	GCCGCTGAA(CGGCGACTT(P L N	CGAGGTGGCC GCTCCACCGG E V A	GAGTGGCGTC CTCACCGCAG E W R R	SCGTCGCCACACCAGCGGGC CGCAGCCGTGTGGTCGCCCG R R H T S G P	4 5 500050 4 5	630
				Bglii				
CGITCTGCCTCCGGGCGTACTCCGGCGGCGCCACTTCTACCTCAACGACCAGTGGCACGAGATCTGCAACGACATCTCCGACCACCTGCTGCTCG GCAAGACGGAGGCCCGCATGAGGGGGGGGGAGAGATGGTGGTGGTCACCGTGCTCTAGACGTTGCTGTAGAGGCTGGTGGACGAGC F C L R A Y S G G H F Y L N D Q W H E I C N D I S D H L L V	GCGGCCACTICI CCGCCGGTGAAGA G H F Y	ACCTCAACGA .TGGAGTTGCT	CCAGTGGCA GGTCACCGT Q W H	CGAGATCTGC SCTCTAGACG	Y TTCTACCTCAACGACCAGGCACGAGATCTGCAACGACATCTCCGACCACCTGCTCG MAAGATGGAGTTGCTGGTCACGTGCTCTAGACGTTGCTGTAGAGGCTGGTGGACGAGC F Y L N D Q W H E I C N D I S D H L V V	CCGACCACC GGCTGGTGG D H L	TGCTCG ACGAGC L V	720

1	1	7	/	164
---	---	---	---	-----

10 20 30 40 50 60 70 80 90 1234567890123456789012345678901234567890123456789012345678901234567890	
Earl ▼ ▼ TGCCCCCGACCACGAGAGAGAGATGGCAGAACCCACGGTGAC TCGGGGGCTGGTCGGAATAGCTTCCTCGCCGCTTCTCTACCGTCTTTGGGTGCCACTG P P T S L I E G A A K R W Q N P R V T	810
ApaI ▼ CGACGACCTGACGGGGGCCCTCACGCACCCCCCCCCCCC	006
D D L T G A L T Q P P L G R T V R A V A D R E L G T H L L E BamHI GACCCGCGGCATCCACTGGCGCGACGCGACCGCTACGCCCACGCGGCGGCGGACGACGCCGTATCCCGCGTA CTGGGCGCCTAGGTGCGGCGCTTGCCGCTGGGCATGCGTGCG	0
TRGIHWIHAANGDPYATVLRGQADDPYPAY BStEII	
CGAGCGGGTGCCTCCCGCGCGCCTCTCTTCAGCCCGACGGGCAGCTACCGCCGATCACGCCCTGGCGGCGAGCATCCTCTG GCTCGCCCACGCACGGGCGCCGCGCGAGGAAGTCGGGCTGCCCGTCGACCCAGTGGCGGCTAGTGCGGGACCGCCGCTCGTAGGAGAC E R V R A R G A L S F S P T G S W V T A D H A L A A S I L C	1080

	1170		1260		1350	٠	1440	
10 20 30 40 50 60 70 80 90 90 12345678901234567890123456789012345678901234567890	CTCGACGGACTTCGGGGGTCTCCGGCGCCGACGGCGTCCCGGTGCCGCAGCAGGTCCTCTCGTACGGGGAGGGCTGTCCGCTGGAGGCGCGCGA GAGCTGCCTGAAGCCCCCAGAGGCCGCGGCTGCCGCAGGGCCACGGCGTCGTCCAGGAGAGGCATGCCCCTCCCGACAGGCGACTCTCGCGCT	STDFGVSGADGVPVPQQVLSYGEGCPLERE	Alwni © GCAGGTGCCGCCGGCGGCGGTGCCGGAGGGCGGGCAGCGTGCCGTGGTCGAGGGGATCCACCGGGAACGCTGGAGGGTCTCGC	Q V L P A A G D V P E G G Q R A V V E G I H R E T L E G L A	GCCGGACCCGTCGCGTCGTACGCCTTCGAGCTGCTGCGGCGGTTTCGTCCGCCCGGCGGTGACGGCCGCTGCCGCCGCCGCTGCTGGTGT CGGCCTGGGCAGCCGCAGCATGCGGAAGCTCGACGACCCGCCAAAGCAGGCGGGCCGCCCACTGCCGGCGGCGGCGGCGGCGGCGGCGACCAAC	PDPSASYAFELLGGFVRPAVTAAAAVLGV	RBrII TCCCGCGGACCGGCGCGCGCATCTGCTGGAGCGGCTCCGGCCGCTGTCCGACAGCCTGCTGGCCGCGCGCG	PADRRADIERLRPISDSLLAPQSLRT

면 되 되

H D

РЕС

G R D

V V L T A

A H

P A G

GRRRL

1	1	9,	/1	64
---	---	----	----	----

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	
Tfil GGTACGGGCGGCGGCGCGCTGGCCGCGCGCTGCTCGCCGGACTCCCCCCGGGGCCCTGCTGTCGGCGCTCGG CCATGCCCGCCGCCGCGCGCGCCCGGACGACGACGACGACGCCCCGGGGCCCCGGGACGAC	1530
BSTEII ©GTCACCGCAGCCGTCCACCGGGAACGCGGTGCTCGCGCTCCTCGCGCATCCCGAGCAGTGGCGGAGCTGTGCGACCGGCCCGG GGTCACCGCAGCCGTCGAGTGGCCCTTGCGCGAGCGGGGGGGG	1620
NOLI GCTCGCGGCGCGCGCGGGGACCCTCCGCTACGACCCGGCTGCTCGACGCCCGGGTGGTCCGCGGGGAGACGGAGCTGGC CGAGCGCCGCCGCCGCCCCCTCTCGGAGCCGATGCTGGCCCCCCCC	1710
Naphi GGGCCGGCGCCGGCCCGGGCCATGTCGTCCTGACCGCGGCCCGGGACCCGGGAGGTCTTCACGGACCCGGAGCGTT CCCGGCCGCCGCCCGGCCCCGCGTACAGCAGGACTGGCGGCGCTTGCCCTGGCCCTCAGAAGTGCCTGGGCCTCGCGAA	1800

120/1	64
-------	----

12345678901234567890123456789012345678901234567890123456789012345678901234567890	
RSTII CGACCTCGCGCCCCCGACGCCGCGCCACCTCGCGCTGCACCCCGGCCGG	1890
DLARPDAAHLALHPAGPYGPVASLVALQA	
GGAGGTCGCGCTGCGGACCCTGGCCGGGCGTTTCCCCGGGCTGCGGCAGGCGGGGGACGTGCTCCGCCCCCCCC	1980
EVALRTLAGRFPGLRQAGDVLRPRAPVGR	
ESPI Bpull02I Alwni Apai Apai CGGGCCGCTCCGGTCCCGGTCCTGAGACACCGGGGCCCCGGTCCGGCCCCCTTCGGACGGA	2070
G P L S V P V S S S	
CACGGGGACGCTCAGACCGTCCCGTGTCCCCGTCCGGCTCCCGTCCGCCCCATCCCGCCCTCCACCGGCAAGGAAGG	2160

FIG. 32F

VDDLVDFA

S

Ω

z

ы

YFHL

L A

G H E

HAL

121/164

CATGCGCGTCCTGCTGACCTCGTTCGCACATCACACTACTACGGCCTGGTGCCCCTGGGCCTGGGCGCCCGCC
Draili GGIGCGGGTCGCCAGCCAGCCTCACGGACACCATCACCGGGTCCGGGCTCCCGGGTGCCGGTCGGCACCGACCACCTCATCCA GGIGCGGGTCGCCCAGCCCAGCCGCGAGCCCTCACCGGCCCCACGCCCACGCCCACGCCGTGGTGGAGTAGGT CCACGCCCAGCCGTCGGTCGCGGAGTGCCTGTGGTAGTGGCCCAGGCCCACGGCCACGGCCACGCCGTGGCTGGTGGAGTAGGT CACGCCCAGCCCA
PvuI ▼ CGAGTACCGGGTGCGGATGGCGGCGCGCGCGCGATCCGGCGTTCGACGAGCCCGTCCCGAGCCGCTGGACTGGGA 2430 GCTCATGGCCCACGCCTACGCCGCGCGGCGTTGGTAGGCGCTTCGAAGCTCCGGCCAGGCTCGGCGACCTGACCT E Y R V R M A G E P N H P A I A F D E A R P E P L D W D
CCACGCCCTCGGCATCCAGGCGATCCTCGCCCCGTACTTCCATCTGCTCGCCAACAACGACTCGATGGTCGACGACCTCGTCGACTTCGC 2520 GGTGCGGGAGCCGTAGCTCCGCTAGGAGGGGCATGAAGGTAGACGAGCGGTTGTTGCTGAGCTACCAGCTGGAGCGAGC

sugar.finalgene b-1 Sequence

FIG. 52G

122/	164
------	-----

		122/164		
	2610	2700	2790	2880
10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	AgeI BSTEII CCGGTCCTGGCAGCCGGACCTGTGGGAGCCGACGACCTACGCGGGCGCCGCCCCAGGTCACCGGTGCCGCGCCCCG GGCCAGGACCTGGACCTCGGCTGCTGGACCTACGCGGCCGCCCCAGGTCACCGGTGCCGCCCCG GGCCAGGACCGTCGGCTGCTGGACCTCGGCTGCTGGATGCGCCCCGCGCGCCGCGCCGCGCCGTCGCCGCCCCGTGCCGCC	Apai GGTCCTGTGGGGGCCCGAÇGTGATGGCCCCGCCGCCGCCGCTGCGGGACCGGCAGCCGCCCGAGCACCGCGCAGCACCCCCC	BetBI AgeI EarI CACCGCGGAGTGGCTGACGTCGACCGCTCCTTCCAAGAGGAGCTGCTCACCGGCCAGTTCACGATCGACCCGAC GTGGCGCCTCACCGACTGCCACTGCCGCCTCCTTCCAAGAGGAGCTGCTCACCGGCCAGTTCACGATCGACCGAC	CCCGCCGAGCCTGCGCCTCGACACGGCCTGCCGACCGTCGGGATGCGTTATGTTCCGTACAACGGCACGTCGTCGTGCCGGACTGGCT GGGCGGCTCGGACGCGGAGCTGTGCCCGGACGGCTGGCAGCCTACGCAATACAAGGCATGTTGCCGTGCAGCCAGGCAGG

INVENTORS NAME: David H. Sherma SERIAL NO.: 09/988,384

REPLACEMENT SHEET

123/164

JACCICGGCAGAGGCGCGICCICGGCGCGCGCGICICGCAGGGGA CTGGGAGCCGCAGAGCGCCACTCCAGGAGCCGCTGCCGCTGCCGCAGAGCGTCCCGCT T L G V S A R E V L G G D G V S Q G D	ICTCGTCGCCACGCTCGAGGTCAGCGCCGAGATCCGCAACTACCCGAAGCA IGAGCAGCGGTGCGAGCTCAGTCGCGCGCTCTAGGCGTTGATGGGCTTCGT L V A T L D A S Q R A E I R N Y P K H	HOLOCTGCCGAGCTCGGCGATCATCCACCACGGGGGGGGGGCGGCTACGCGACCGGGCGGCGCTGGGTGGTGCCGCCCCCCCGCCCG	CGCCGAGCTGTGGGACGCGCGGTCAAGGCGCGGGGCGGCCGTCGCCGAGCAGGGGGGGG
CCICGGCGCCGCCGCGAGGGGGGCGCGCGCGCGCGCGCGC	CGCCGAGATCCGCAACTAC GCGGCTCTAGGCGTTGATG A E I R N Y	CACGGCGGGCGCCACC GTGCCGCCCCGTGG H G G A G T	GGCCGTCGCCGAGCAG CCGGCAGCGGCTCGTC A V A E Q
CCICGGCGGCGACGGGGGGGGGGGGGGGGGGGGGGGGGG	CCCCCAGATC CCCCCTCTAG	CACGGCGGG	GGCCGTC CCGGCAG
GGAG 1	ប្តី ថ្មី 🗎	· ·	20 CG
CTCCA E V	AGTCAGCG TCAGTCGC S Q R	ATCATCCAC TAGTAGGTG I I H	GTCAAGGCG CAGTTCCGC V K A
iccececei Aggcgcgca S A R	CTCGACGCG GAGCTGCGC L D A	TGCTCGGCG	GACGCGCCG CTGCGCGGC D A P
CICGCCGCAG	GTCGCCACG CAGCGGTGC V A T	CTGCCGAGC GACGGCTCG L P S	GAGCTGTGG SCTCGACACC E L W
ACGGACTGG	ATCGAGCTC TAGCTCGAG I E L	CACGCGCTC GTGCGCGAG H A L	ATGCTCGCC TACGAGCGC M L A
CCCGGGCCAG	SACCTCGAC CTGGAGCTG D L D	STGCCGATG CACGGCTAC V P M	ccccaccic seccrccac P Q V
CCGCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CGCTCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CGGACTTCCCGCTGAAGC	ACGCGGTG
AGIGAGCCGC FCACTCGGCG	ATCCTGGAGG	ACCCGGTTCA FGGGCCAAGI F R F I	CGCCGTGATCAACGCGGTGCTGCTGCT GCGGCACTAGTTGCGCCACGGCGTCCAGTACGA A V I N A V P Q V M L
	CHCACTCGGCGGCGCCGGGGCCCAGACGGAGCCGCAGAGCGCGCACTCCAGGGCGCCGCCGCTGCCGCAGAGCGTCCCGCT SEPPARC CGGGGCCCGGGGCCCAGACGGAGCCGCAGAGCGCGCAGAGCGTCCCGCT SEPPARC GGGGCCCGGGGGCCCAGACGGAGCCGCAGAGCGTCCCGCTGCCGCAGAGCGTCCCGCT SEPPARC GGGGCCCAGACGGACTGGGAGCCGCAGAGCGCTCCCGCAGAGCGTCCCGCTGCCGCAGAGCGTCCCGCTGCCGCAGAGCGTCCCGCTGCCGCAGAGCGTCCCGCTGCCGCAGAGCGTCCCGCTGCCGCAGAGCGTCCCGCTGCCGCAGAGCGTCCCGCTGCCGCAGAGCGTCCCGCTGCCGCAGAGCGCTCCCGCTGCCGCAGAGCGCTCCCGCTGCAGAGCGCTCCCGCTGCAGAGCGCTCCCGCTGCAGAGCGCTCCGCTGCAGAGCGCGCCGCTCCGCTGCAGAGCGCGCCGCTGCAGAGCGCGCCGCTGCAGAGCGCGCCGCTGCAGAGCGCGCCGCTGCAGAGCGCGCCTGCAGAGCGCGCTCCGCTGCAGAGCGCGCCGCTGCAGAGCGCGCCGCTGCAGAGCGCGCCTGCTGCAGAGCGCGCCGCTGCAGAGCGCGCCTGCAGAGCGCCGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCAGAGCCGCTGCAGAGCCGCTGCAGAGCCGCTGCAGAGCCGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCAGAGCCGCTGCAGAGCCGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCTGCAGAGCCGCCTGCTGCAGAGCCGCTGCTGCAGAGCCGCTGCAGAGCCCGCTGCAGAGCCGCTGCAGAGCCCGCTGCAGAGCCCGCTGCAGAGCCGCCTGCAGAGCCCGCTGCAGAGCCCGCTGCAGAGCCCGCTGCAGAGCCCGCCTGCAGAGCCCGCTGCAGAGCCCGCTGCAGAGCCCGCTGCAGAGCCCGCTGCAGAGCCCGCCTGCAGAGCCCGCCTGCAGAGCCCCCCTGCCAGAGCCCCCTGCCAGAGCCCCCCCC	CATCCTGGAGGCGCTCGACTCGACTCGCCACGCGCGCCCCCCCGCGCGCCGCTCCCGCTGCCGCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCGCCCCTCCGCCG	CATCTGGAGGCGCTCGCCGACCTCGACATCGAGGCGCCCTCGAGGCGCCCCTCCGCAGGCGCCCCCCGCTCCTC

sugar.finalgene b-1 Sequence

FIG. 32I

3600

124/164

06 06	3330	3420	3510 T
10 20 30 80 90 90 50 1234567890123456789012345678901234567890123456789012345678901234567890	GTICTICCIGCCGCCGAGCTCACGCCGCAGGCCGTGCGGACGCCGTCGTCGCATCCTCGACGACCCTCGGTCGCCACCGCCGC CAAGAAGGACGGCGGCGGCGTCGGGCGTCCGGCACGCCTGCGGCGTAGGAGCTGCTGGGGAGCCAGCGGTGGCGGCG F F L P P A B L T P Q A V R D A V R I L D D P S V A T A A	GCACCGGCTGCGCGAGGAGCTTTCGGCGACCCCACCCCGGCATCGTCCCCGAGCTGGAGCGGCTCGCCGCGCGCAGCAGCGCCCCCCCC	StuI GCCGGCCGACGCCCGCACCCCACCCCTCGCCCCAGGCCTCACCCCTGTATCTGCGCCGGGGGACGCCCCGGGCCCACCCTCCGA CGGCCGGCTGCGGGGTCCGGAGTGGGGACATAGACGCGGCCCTTGCGGGGGGGG
70 23456789012345	ACCGIGCGGGACGCCGICGICCGCATCCICGACGACCCTCGGICGCCACCGCCCCCCACCGCAGCCAGCGAGCCAGCGGAGCCAGCGGGAGCCAGCGGGGGAGCCAGCGGGGGG	AGCGGCTCGCCGC STCGCCGAGCGGCG SRLAAA	CGGGGGACGCCC
60 3456789012	GTCCGCATCC CAGGCGTAGC V R I I	CCCGAGCTGC GGGCTCGACC P E L E	GTATCTGCGC
50 3456789012	SGACGCCGTC CCTGCGGCAG D A V	GGGATCGTC SCCTAGCAG G I V	StuI ▼ HCTCACCCT
40 3456789012:	AGGCCGTGCGC TCCGGCACGCC	CCACCCGGC GGTGGGGCCGG T P A	TCGCCCCAGG
30 3456789012	CTCACGCCGC GAGTGCGGCG L T P Q	TTCGGCGACC AAGCCGCTGG F G D P	GCCGCACCCC
20 3456789012	SCCGGCCGAG CGGCCGGCTC	CGAGGAGACC SCTCCTCTGG E E T	CCGGCAC <u>IGA</u> GGCCGTGACT R H
10 123456789012:	STICTICCIGCCGGCCGAGCICACGCCGC SAGAAGGACGGCGGCCGGCTCGAGIGCGGCG F F L P P A E L I P Q	SCACCGGCTGCGCGAGGAGCCTTCGGCGGCGACCC SGTGGCCGACGCGCTCCTGGAAGCCGCTGGC H R L R E E T F G D P	SCCGGCCGACGCCCGGCF SGCCGGCTGCGGGCCGT

sugar.finalgene b-1 Sequence

AAGACCGAAAGGAGGAGCACCGIGIACGAAGTCGACCACGCCGACGTCTACGACCTCTTCTACCGGTCGGGCAAGGACTACGCCGCC TTCTGGCTTTCGTCCTCGTGGCACATGCTTCAGCTGGTGCGGCTGCAGATGCTGGAGAAGATGGACCCAGCGCCGTTCCTGATGCGGCGG K, Z, Ö pz; Ö Earl Λ Ω 耳 > 凶

FIG. 32J

REPLACEMENT SH

3960

× Ы

G > ഗ ഗ Ŀ Σ വ

æ Ø ſz, ×

足 Ö Ч

Ŀ A 姳 Σ Ω

Q Ø

耳 Ц H

BbsI

	3690	3780	3870
90 34567890	CGCATCTG SCGTAGAC H L	CCGACGCC 3GCTGCGG D A	ACCTGAAG IGGACTTC
80 56789012	TACGGGCA ATGCCCGT	GCGGCTGC CGCCGACG R L P	CGTCGGCT
70 78901234	GCCTGCGG CGGACGCC A C G	GCCCGCAA CGGGCGTT A R K	Nsphi ▼ TTCAGCTC
10 20 30 40 50 50 90 90 90 234567890123456789012345678901234567890123456789012345678901234567890	StuI ▼ GAGGCCTCCGACTGCCGACCTCCCGTACCCCCGAGGCCTCCTCGCTCG	Styl GAGCACTTCACCAAGGAGTTCGGCGACACCGCCGGCTGGTCCGAGGACATGCTCACCCCACGCCCGCAAGCGGCTGCCCGACGCC CTCGTGAAGTGGTTCCTCAAGCCGCTGTGGCGGCCTCGCCGACGGCTGCCGGCCG	NSPHI ACGCTCCACCAGGGCGACATGCGGGCTCGGCCGGAAGTTCTCCGCCGTGGTCAGCATGTTCAGCTCCGTCGGCTACCTGAAG TGCGAGGTGGTCCCGCTGTACCTGGCCGAAGCTTCTAGCTCCAGGTCAGCTCAAGTCGAGGCGGCAAG
0 012345670	I CCTCGCTC GGAGCGAG	AGGACATG	CCGCCGTG
50 234567890	StuI * CGAGGCCTC	AGCTGTCCG CGACAGGC L S E	gaagttet etteaaga
40 345678901	CCGTACCCC GCATGGGG	CGGCCTGG2 SCCGGACCT G L E	SCTCGGCCC CGAGCCGGC
30 56789012:	GTGCGCTC CACGCGAG(V R S	GACACCGC CTGTGGCG D T A	HI GACTTCCG CTGAAGGC
20 78901234	CCGACCTG GCCTGGAC	AGTTCGGC TCAAGCCG	Nsphi ** ** ** ** ** ** ** ** ** ** ** ** *
10	StuI ▼ GCCTCCGACATCGCCGACCTGGTGCGCT CGGAGGCTGTAGCGGCTGGACCACGCGA ASDIADLVRS	Styi GAGCACTTCACCAAGGAGTTCGGCGACACCG CTCGTGAAGTGGTTCCTCAAGCCGCTGTGGC E H F T K E F G D T A	ACCAGGGGG GGTCCCGC
12345678	StuI GAGGCCTC CTCCGGAGG E A S	GAGCACT CTCGTGA! E H F	ACGCTCC

sugar.finalgene b-1 Sequence

ACGACCGAGGAACTCGGCGCGGCCGTCGCTTCGCGGAGCACCTGGAGCCCCGGTGGCGTCGTCGTCGTCGAGCCGTGGTGGTTCCCCG TGCTGGCTCCTTGAGCCGCCCCCGGCGAGCAAGCGCCTCGTGGACCTCGGGCCACCGCAGCAGCAGCAGCACCACCACAAGGGC
T T E E L G A A V A S F A E H L E P G G V V V V E P W W F P

126/164

10 20 30 40 50 60 70 80 90 123456789012345678901234567890123456789012345678901234567890 Aatii
GAGACCTICGCCGACGTCAGCGCCGACGICGTCGCCGTGACGGGCCACCGTGGCCCGTGTCTCGCACTCGGTGCGGGGGGGG
AACGCGACGCGATGGAGGTCCACTTCACCGTGGCCGAAGGGCGTGCGGCACTTCTCCGACGTCCATCTCATCACCTGTTC 4140 TTGCGCTGCGCGCTACCTCCAGGTGAAGTGGCACCGGTGCGCGCGC
Sfil Earl CACCAGGCCGAGTACGAGGCCGCGCTCACGGGCTCGAGTACCTGGAGGGCGGCCGTCGGGCCGTGGCCTTCGTC GTGGTCCGGCTCATGCTCCGGCGCGCGCGCGCGCGCGCGC
Aatii Apali GGCGTCCCCCCCTGAGCACCCCCCCGGGGGGGGGGGGACGTCCCGGGTGCCAAGCAAAGAGAGAAAACGAAAAAAAA

FIG. 32L

4410 4500 12345678901234567890123456789012345678901234567890123456789012345678901234567890 GAGAAGATCAGCTTCGTCCACTGGGCGCTGGACCCCGACCGGCAGGACGTCGGCTACCTTCCCGGCGTGCCGCGTTCTGGGCATCCCGGAG CTCTTCTAGTCGAAGCAGGTGACCCGCGACCTGGGGCTGGCCGTCTTGCAGCCGATGGAAGGGCCGCACGCGCAGACCCGTAGGGCCTTC <u>AAGACCCGAATACCGCGTGTCCGCCGCGGCCGCACCACGCCCTAGGGCCTTCACCCTGGCCGTCGTCGCACCCTGCTGGCGGGCACCGGCACCACCACC</u> GTGGCGGCCGCCCCCGGCCCCCGCCGACACGCCAATGTTCAGTACACGAGCCGGGCGGCGGAGCTCGTCGCCCAGATGACGCTCGAC TICTGGGCTTATGGCGCACAGGCGCCCCGGCGTGCGGGTCCCGGAAGTGGGACCGGCAGCAGCAGCGTGGGACGACGACGCCGTGGTGG ט Д Ö Ŋ Д ᄓ ဗ z ø ø æ œ Д Ω Д ద Not I ᆸ O ø æ 3 24 Ħ > 24

sugar.finalgene b-1 Sequence

FIG. 32M

ø

Ö

П

æ

Ŋ

Z

CTGCGTGCCGCCCGACGGCCCCGAACGGCATCCGCCTGGTGGGGCAGACCGCCACCGCGCGCTGCCCGCGGCCGGTCGCCCTGGCCAGCACCTTC

MscI Ball

1	28	/1	64	
---	----	----	----	--

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	
Styl Ncol Ncol T GACGACACCATGGCCGACAGGTCATGGGCCGCGACGGTCGCGCGCTCAACCAGGACATGGTCCTGGGCCCGATGATGAAC CTGCTGTGGTACGGCTGATGGTTCAGTACCGGCGTACTTG CTGCTGTGGTACCGGCTGTCGATGGCTACCGGCGCTACTTG CTGCTGTGGTACCGGCTGTCGATGCCGTTCCAGTACCTACTTG D D T M A D S Y G K ·V M G R D G R A L N Q D M V L G P M M N	4770
AACATCCGGGTGCCGCACGGCGGCCGGAACTACGAGGACCTTCAGCGCCCCCCCC	4860
GGCATCCAGGGTGTGGAGCACGGCCAAGCACTTCGCGGCCAACAACCAGGAGAACAACCGCTTCTCCGTGAACGACCATGTC CCGTAGGTCCCACGCCCAGACTACTGGTGCCGGTTCGTGGTCCTTGTTGGCGAAGAGAGAACAGAGCACTTGCGGTTACAG G I Q G A G L M T T A K H F A A N N Q E N N R F S V N A N V	4950
Styl Sfil GACGAGCAGACGCTCCGCGAGATCGAGTTCCCGGCGTTCGAGGCGTCCTCCAAGGCCGGCGCGCGC	5040

FIG. 52N

129/164

10 20 30 40 50 60 70 80 90 123456789012345678901234567890123456789012345678901234567890		ı
FSPI Draiii T CTCAACGGGAAGCCGTCCTGCGGCGAGCTCCTCAACAACGTGCTGCGACGCCAGGGCTTCCAGGGCTGGGTGATGTCCGAC GAGTTGCCCTTCGGCAGGACTTGTTGCACGACGCGTGACGCCCGAAGGTCCCGACCCAACACTACAGGCTG L N G K P S C G N D E L L N N V L R T Q W G F Q G W V M S D	5130	
Styl EcoNI PflMI TGGCTCGCCACCCGGCCATCACCAAGGGCCTCGACCTCCCCGGCGACGTCCCGAAGGGCGAGA ACCGAGCGTGGGCGTGCGTAGTGGTTCCCGGAGGTCGTCCTCTCCCGCGGGCGCTGCTGCTCCCGGAGGCTTCCCGCTC W L A T P G T D A I T K G L D Q E M G V E L P G D V P K G E	5220	129/104
BbsI ▼ CCTCGCCGCCGGCCAAGTTCTTCGGCGAGGCGCTTGAAGGCGGCACGGTCCCCGAGGCGGCCGTGACGCGGCG GGGAGCGGCGGCGGCGGCGGCTTCTGCCGGCAGGCGCAGGCGCCGCCGCCGCCGCCGCCGCCGCC	5310	
GAGCGGATCGTCGGCCAGATGGAGAAGTTCGGTCTGCTCCTCGCCACCCGGCGCCCCGAGCGCGCGACAAGGCGGGTGCCCAGGCG CTCGCCTAGCAGCCGGTCTACCTCTTCAAGCCAGAGGAGGGGGGGCGCGGGGCGCGGGGCTCGCGGTTCCGCCCACGGGTCCGC E R I V G Q M E K F G L L A T P A P R P E R D K A G A Q A	5400	

INVENTORS NAME: David H. Sherman et al. SERIAL NO.: 09/988,384

REPLACEMENT SHEET

1	30		1	64	
---	----	--	---	----	--

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901	
FSPI Alwni GTGTCCCGCAAGGTCGCCGAAGGGCGCGGTGCTCCTGCGCAAGGGCCAAGGGCCAAGGGCCAAGGCATC CACAGGGCGTTCCAGCGGCTTTGCCGCCACGAGGACGCGTTGCTCCGGGACGCCGACGCCGTTCTCGTAG V S R K V A E N G A V L L R N E G Q A L P L A G D A G K S I	5490
Styl GCGGTCATCGGCCCCGACGCCCTCGACCCCAAGGGCCTTGGGCACTCCTCGGCGCGCGC	5580
ACCATCAAGGCCCGCGCGGGGGGGGGGGGGGGGGGGGGG	5670
xhoI PaeR7I ▼ AGCCCGGCGTTCAACCAGGCCACCCGGGCAAGGCGGGGGGGG	5760

REPLACEMENT SHEET

Д z

Ω

ഗ

Н

× Д

E-1 Ö

ρι

Ø Н Z Д ĸ Ω.

Ö

囟

H Q Ω Д

1	3	1/	1	64	

10 20 30 40 50 60 70 80 90	
123456789012345678	
RSFII AgeI TACCGCATCGCGGTCCGTGCTTACGCCACGGTGCAGCTCGGCAGCCCATCGAGGCCGGTCAGGTCTACGGCAAGGTG ATGGCGTAGCGCCACGGTGCCACGTGCCACGTCGAGCCGTCGGTGGTGGTGGTGGTGGTGGTGGCGTGCCAGTCCAGATGCCGTTCCACATCCAGATGCCGTTCCACAGTCCAGATGCCGTTCCACAGTCCAGATGCCGTTCCAGATGCCGTTCCAGATGCCGTTCCAGATGCCGTTCCAGATGCCGTTCCAGATGCCGTTCCAGATGCCGTTCCAGATGCCGTTCCACAGTCCAGATGCCGTTCCACAGATGCCGTTCCACAGTCCAGATGCCGTTCCACAGTCCAGATGCCACAGATGCCACAGATGCCACAGATGCCACAATGCCACAATGCCACACACA	5850
StyI AGCAGCCCGCTCCTCAAGCTGACGCACGCACCGATCTCGGGCTTCGCGATGAGTGCCACCCCGCTCTCCCTGGAGCTG TCGTCGGGCGAGGAGTTCGACTGCTGTTCGAGTGCTAGGCCTACTCACGGTGGGGCGAGGAGCTCGAC S S P L L K L T K G T H K L T I S G F A M S A T P L S L E L	5940
NruI PvuI ▼ GGCTGGGTGACGCCGGCGGCGGCGACGACGCGAAGGCCGTGGAAGGCCCGTACGCCGGTCGTCTTCGCCTAC CCGACCCACTGCGCCGCCGGCTGCTTCCGGCACCTCCAGCCGCCCTTCCGGCGCTTCCGGCGATGCCGCCAGCAGCGCATG G W V T P A A D A T I A K A V E S A R K A R T A V V F A Y	6030
GACGACGGCACCGAGGGCGTCGACCGTCCGAACCTGTCGCTGCCGGGTACGCAGGACAAGCTGATCTCGGCTGTCGCGGACGCAACCCG	6120

sugar.finalgene b-1 Sequence

CTGCTGCCGTGGCTCCCGCAGCTGGCCTTGGACAGCGACGGCCCCATGCGTCCTGTTCGACTAGAGCCGACAGCGCCTGCGGTTGGGC

6480

132/164

6300 8 CAGGCGCGCCGAGGCCACCGCCGCGCGCTGCTCTACGGTGACGTCAACCCGAGCGGCAAGCTCACGCAGAGCTTCCCGGCCGCCGAGAAC CAGCACGCGGTCGCCGGCGACCCCGACAAGCTACCCGGGCGTCGACAACCAGCAGAGGTACCGCGAGGGCATCCACGTCGGGTACCGCTGG GTCGTGCGCCAGCGGCCGCTGGTCGATGGGCCCGCAGCTGTTGGTCGTCTGCATGGCGCTCCCGTAGGTGCAGCCCATGGCGACC 123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890 GICCGCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCAGTIGGGCTCGCCGTTCGAGTGCGTCTCGAAGGGCCCGCGCCTTTG <u>AACACGATCGTGGTCCTCAACACCGGGTTCGTCGGTGCTGATGCCGTGGCTGTCCAAGACCCGCGCGGTCCTGGACATGTGGTACCCGGGC</u> Acc651 NepHI Afliii 8 Σ П П > Ø 8 ტ ¥ Н 3 А 各 ט ິທ AgeI Ö 2 z

sugar.finalgene b-1 Sequence

AAGCTGTTCCTCTTGCAGTTCGGCGACAAGGGCAAGCCCGTGCCGGACAGGCATGTGGAGCAAGTGCGTCTCGCGGGGCTGGCAGCACCAC TTCGACAAGGAGAACGTCAAGCCGCTGTTCCCGGTTCGGGCACGGCCTGTCGTACACCTCGTTCACGCAGAGGCGCCCCCGACCGTCGTGCGT ט 闰 24 Œ ഗ ⊱ Ø ø Z Д ෆ Ö Д Ö H O ᄺ

IG. 32R

133/164

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901	
RSTII ACGTCCACGGGTGGTCTGAAGGTCACGGTCACGGTCGTCCGAAGCGCGCCGGCCAGGAGGTCGTCCAGGCGTACCTCGGTGCC TGCAGGTGCCCACCAGACTTCCAGTGCCAGGCGTTGTCGCCCTTCGCGCGGCCGGTCCTCCAGGCGTCCGCATGGAGCCACGG TGCAGGTGCCCACCAGACTTCCAGTGCCAGGCGTTGTCGCCCTTCGCGCGCCGGTCCTCCAGCAGGTCCGCATGGAGCCACGG T S T G G L K V T V T N S G K R A G Q E V V Q A Y L G A	6570
BBSI AGCCCGAACGTGACGCTCCGCAGGCGAAGAAGCTCGTGGGCTACACGAAGGTCTCGCTCG	0999
SfcI PvuII AgeI ▼▼ ▼▼ GTGAACGTCGACGCCACGGTCTGGGGACCTGCGGGGCAGCGCCACGGTCTGGTGACGTGACGTGACGCCCCCCCACGGTCAACGTCTGGTGACGTGACGCCCCCCCC	6750
Stul Mael Bfal Bfal GTGAAAGCGGCGGCGCCACCCGCACCCGCCACCCGCTTTTTCGCCTGCTGGGCCTGGTGGCCTGGTTGGCCTGGTTGGCCTGGTCGGACCACCTGGACTAGGCCGGACCACCTGGACTAGGCCTGGACTAGGCCTGGACTAGGCCTGGACTAGGCCTGGACTAGGCCTGGACTAGGACCACGACCACGACCACGACCA	6840

FIG. 32S

INVENTORS NAME: David H. Sherman et al.

REPLACEMENT SHEET

7200

SERIAL NO.: 09/988,384 134/164 7020 GCCGCGCGAAGCTCTCGGCCCCGCGGGGGGGGGCCTTCCGGCGGTGCCTCGCCCGCGTAGGCGGGGGAGAGGTGCACGGGTACCGGGTAGT CGGCGCGCTTCGAGAGCCGGGCGCCCTCGCAAGGCCGCCACGGAGCGGGCGCATCCGCCCCGCTCTCCACGTGCCCATGGCCCATCA GCTGGGCGAGCCGGGTAAGCGCGCGCGCAGCTAGTGGGCGTCGCGGACGCCCGCGAGGTCCACGCCCGGCTAGCCGTCCGACTCCTGGA CGACCCGCTCGGCCCCATTCGCGCACGGCGTCGATCACCCGCAGCGCCTTGCGGGCGCTCCCAGGTGCGGGCCGATCGGCAGGCTGAGGACCT Apali Acc651 ഗ Ö Earl တ Д 团 æ ტ

123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890

Alwni

ᆸ

R

臼

sugar.finalgene b-1 Sequence

8

8

20

PvuI

GCGTGAGGGTGTCGATGCCGCGGGGGTCGAGGTGGCTGCGCAGCTCGTCGCGGCGCTCGGTGCGCACGGTGAAGAGAGGTGCCAGACCGGGT CGCACTCCCACAGCTACGGCGCCCGCAGCTCCACCGACGCGTCGAGCAGCGGCGAGCCACGTGCCACTTCTCCACGGTCTGGCCCA FapI

囟

Д

Ö

回

7110

Ħ

H

24

H

臼

跘

œ

Д

M

Н

ഗ

出

Н

ρ

æ

ט

Ħ

BatEII

CGGTGTCGGGCGCGGTCACCGGCAGGCCGATGCCGGGCAGTCCGGCGAGCCCCGGAGAGGTACTCCGCGGCCAGCGCCGACCTGCGGCCGT ď 臼 ഗ æ ტ Ö ט ы Дí

135/164

10 20 30 40 50 60 70 80 90 12345678901234567890123456789012345678901234567890123456789012345678901234567890	30 234567890123456	40 50 78901234567890123	60 456789012345678	70 80 9012345678901234567	90 7890	
PVULI PVULI TCCAGCTGTCCAGGTGGGCGAGCCGGAGCACGGCGGCGTGCATCTCGTCCAGGCGGGAGTTGGTGCCCTTCGTCTCGTGGTGGCTGT AGGTCGACAGGTCCACGCTCGGCTAGGCGTCGTCGTCGCGGAGGTTGGTGCCCTTCGTCGTGGTGCCACAAGACAAGAAGAAAAAAAA	BamHI ▼ AGCCGGATCCGCAGCP FCGGCTAGGCTCGT	XcmI ▼ GCACGGCGTGCATCTCGTCCAGGCGGGAGTTGGTGCCCTTCGTCTCGT CGTGCCGCCGACGTAGAGCAGGTCGCCCTCAACCACGGGAAGCAGAAGCAGTCGCCCTCAACCACGGGAAGCAGAAGCAGAAGCAGAAGCAGAAGCAGAAGCAAGCAAAGCAAAGCAAAGCAAAGCAAAGCAAAAAA	XcmI * ** ** ** ** ** ** ** ** ** ** ** **	GTGCCCTTCGTCTCGTGGC CACGGGAAGCAGAGCACCC T G K T E H		7290
BSPEI ACTICIGCCGCGAGCCGTAGTTGCGGAGCAGCCGTTCGGCGAGCCCCCCCGCCGCCGCCGCCGCCGCCGCCGCCGC	BSPMII BSPEI TGCGGAGCATCCGGAG	PMII PEI GGAGCCGTTCGGCGAGCTCGG CCTCGGCAAGCCGCTCGAGCC	eggregeeggegeegeegeegeegeegeegeegeegeege	ACGGCGCCGCCGTCGCCG?		7380
AGCCGAGGTTCTTGCCCCGGTAGAAGGTGAACGCGGCCGACGACCCGGCGCCGGCCCGGCCCCGGGTAGCGGGCGCCGTGGGCCTT TCGGCTCCAAGAACGGGCCCATCTTTGCGCCGGTGGCTGCTGGGCCGCGGGCTAGGCGGCCGGGGCCATCGCCCGCGGGCACCCGGA C G L N K G P Y F S F A A V S S G A G I R R G R Y R A G H A	TAGAAGCTGAACGCGC ATCTTCGACTTGCGCC Y F S F A	GCCACCGACGACCGGCG CGGTGGCTGCGCCGC A V S S G A	BACCCGGCGCCGATCCGCCGGCCCCGGTAGCGGC STGGGCCGCGGCTAGGCGGCCGATCGCCC S G A G I R R G R Y R	CGGTAGCGGGCGCCGTGG(GCCATCGCCCGCGGCACC(R Y R A G H		7470
RSTII GCGCGGCGTCCTCGACGATGTGCAGGCCGTGCCGGAGCTCGCGGAGGCGTCCATGTCGGCGGGGTGCCCGTAGAGGTGGACGG CGCGCCGCAGGAGCTGCTGCCGGCCAGGCGCTCGAGCGCCTCCGCAGGTACAGCGCCCCACGGGCATCTCCACCTGCC QAAD BVIHL GHRDALL BRLBNL BVU	Re ▼ GCAGGCCGTGCCGG ACGTCCGGCACGGCC H L G H R	RsrII ▼ ∋GTCCGCGAGCTCGCGGAGC CCAGGCGCTCGAGCGCCTCC R D A L E R L	GCGTCCATGTCGGCC CCGCAGGTACAGCCGC	RSTII▼ ▼ GGTCCGCGAGCTCGCGGAGGCGTCCATGTCGGCGGGGGGGCCCGTAGAGGTGGACG CCAGGCGCTCCAGGCAGGTACAGCCGCCCCACGGGCATCTCCACCTG		7560

FIG. 32U

7920

10 20	30 40	50	09	70	80	06	
123456789012345678	0123456789012	34567890123456	78901234567	8901234567	8901234567	890	
GGAGGAGCGCCCGGGTGCGGGTGATCGCCTTCTCGACGAGCGGGGTCCAGGGTGGGT	GCCTTCTCGACGA GGGAAGAGCTGCT A K E V	TTCTCGACGAGCGGGTCCAGGGTGGGTGCTCCTCGTGCGGCTCGACGGGCACC AAGAGCTGCTCGTCGCCCAGGACCACCGAGCAGCCGGGCCGTGC KEVLLPPDEYP	GTGGGGTGGTC CACCCCACCAG T P H I	CTCGTGCGGC GAGCACGCCC	TCGACGGCA PAGCTGCCCGT E V P		7650
BSAAI PWLI AFILIII GGGTCGCGCCGGTGGCGACCACCGCGACGACGATCCCCCGGGTCCGATGCCGAGGC CCCAGCGCCACCGCCTCGTCGCGACGACGATGCTCCC CCCAGCGCCGCCCTCGTCGCGACGACGACGCTACTCCCCTCCTCCTCCTCCCAGGCCCAGGCTACGGCTCCCGATGCCAGGCTCCCC PTAGTA GTAS VALWSAIVATACATGCACCCCTCCTACTGCAGGGCCCAGGCTACGGCTCCCC	Pvuli ▼ CAGCTGGCGATG1 GGTCGACCGCTAC2	BsaAI Aflii ▼ TGGCGATGTGCGAGGGGACGATCACCTCGTCCCCGGGTCCGATGCCGAGG ACCGCTACATGCACGCTCCCTGCTAGGGGGCCCAGGCTACGCTACAGGGCCCAGGCTACAGGGCCCAGGCTACAGGGCCCAGGCTACAGGGCCCAGGCTACAGGCCTACAGGGCCCAGGCTACAGGCCTACAGGCCCAGGCTACAGGCCTACAGGCCCAGGCTACAGGCCTACAGGCCCAGGCTACAGGCCCAGGCTACAGGCCCAGGCTACAGGCCCAGGCTACAGGCCCAGGCTACAGGCCCAGGCTACAGGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCTACAGAGCAGCCCAGGCTACAGAGCAGCCTACAGAGCAGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCCCAGGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCAGCAGCCCAGGCTACAGAGCAGCCCAGGCTACAGAGCAGCAGCAGCCAGAGCAGCAGCAGCAGCAGCAGC	ACGATCACCTC TGCTAGTGGAG	GTCCCGGGG SCAGGGGCCC S D G P	CCGATGCCGA AGGCTACGGCT G I G		7740
Tfil CGCGGAGGCCAGCTGGAGGCGTCCATCCCGCTGTTCACGCCGACGGCGTGGTCCGTCTCGCAGTACGCGGCGAACTCCGCCTCGAATC GCGCCTCCCGCTCGAGGTAGGGCGACAAGTGCGGCTGCCGCAGCAGCAGTACGCGCCGCTTGAGGCGAGCTTAG G R L A L Q L A D M G S N V G V A H D T E C Y A A F E R E F	CCGCTGTTCACGC SGGCGACAAGTGCC	Tfil TGTTCACGCCGACGCGTGGTCCGTCTCGCAGTACGCGGCGAACTCCGCCTCGAATC GACAAGTGCGGCTGCCGCCAGAGCGTCTCGCAGTACGCGCCGCTTGAGGCGGAGCTTAG S N V G V A H D T E C Y A A F E A E F	GTCTCGCAGT? CAGAGCGTCA1 T E C)	ACGCGGCGAA(GGCGCCGCTTC AAF	T TCCGCCTCGA		7830
		PvuI					

sugar.finalgene b-1 Sequence

CTTCGAGTTCGGGTCCGAGGTAGCGCCCCCGAGTCGAGGACGCGGCGATCGCGGCGTCGGTCTCCGCGCGGAGCTCCTCGTAGGCGG

ט

Q

R D

囟

Ç

Ö

 Ω

闰

LERN

E E

r r

SD

A G L

137/164	
---------	--

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	
BSTXI PVUII CCTTGAGGTCGAGGGACGCGGGGGGTCTCGGCGGCGGCTGCTCACGGGGGGGG	8010
5 5 F	8100
BEDEI BEGEI GCCGAGCTCGCGCTCGATCTCGCCGTCGAGGAGTAGCGCAGGTCGTGGCCCTTGCGGTCGGCGACCTTCCGGACCGAGGACCAGTC CGGCTCGAGCGCGCGGCGGCGGCGCCTCCTCATCGCGTCGGCGACCGAGGCCTCGTCGGTCAGC G L E R E I E G G D L S Y R L D H G K R D A V K R V S S W D	8190
Bglii GGCGCCGAGCGAGTCCAGGAGGATGCCGGTGAGTTCGCGGTTGGTCAGCTCCAGGCCGCCGCCGATGTGGTAGATCTCGCCGGCCCGGCCCCGCCCCGCCCG	8280

REPLACEMENT SHEET

Ø

Ω

Ç

>

Ö

2

>

>

;

>

₽

Ω

闰

>

×

Ö

ß

н

Д

വ

Ö

ഗ

3

H

囟

വ

ល

8640

138/164

8550 8370 8460 Н CCCGCCGTCGAGGAGGTTCGTCACGAAGAGGGGGATGAGCTTCTCGGGGTGCTGGTACGGCCCGTAGTTGTTGCAGCAGCGGGTGATCCG പ്പ TACGTCGAGGCCGTACGTCCGGTGGTAGGCGGCGAACGAGGTCGGAGCCGGCCTTGGACGCCGCGTAGGGCGAGTTGGGCTCCAGCG **ATGCAGCTCCGGCATGCAGGCCACCATCCGCGCCCGTTGCTCCAGCCTCGGCCGGAACCTGCGGCGCATCCCGCTCAACCCGAGGTCGCC** GCCCGCGAGGACGAGCGCGATGCCCCCGGCAGTGGTCGTCGGTGTGCACCCACTCGCGGACGTTCGCGCCGTCGCCGTACAGCGGAGCGT **GGGCGGCAGCTCCTCCAAGCAGTGCTTCTCCCCCTACTCGAAGAGCCCCACGACCATGCCGGGCATCAACAACGTCGTCGCCCACTAGGC** CGGGCGCTCCTGCTCGCGCTACGGGGCCGTCACCAGCAGCCACACGTGGGTGAGCGCCCTGCAAGCGCGGCAGCGGCATGTCGCCTCGCA 1234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890Д Н Д H Ы O × U Ö z Д Z ט × Ø O Z Д > Styl 足 × α 臼 3 耳 Apali > Д Ħ 闰 × H Н Д А Н 二 Д Ы Ö G4 R U > Earl H Н z ø H H Ы > Ω Ы O ø Ö O

sugar.finalgene b-1 Sequence

Д Ы 囟 Д Z တ Д × ď ApaLI 4 ഗ × ď Ö S Ω ᆸ > ø PvuI ĸ ø Ħ æ H × Ö П Ω >

GCTGCTCTCGGTCCAGGAGCCGGAGTCGATCGACCCGTACACCTCGTCGGTGGAGACGTGCACGACCCGGCCGACGCCGGCGGCGTCGACGGC CGACGAGAGCCAGGTCCTCGGCCTCAGCTGGGCATGTGGAGCAGCCACCTCTGCACGTGCTGGGCCGGCTGCGGCCGCTGCACGTGCTGCCG

0006

Bbsi

BSTEII

139/164

8820 8910 8730 А CGCGTCCACCGGGGGGAGGTTGGCGGGGTTGCCCGCGTAGGTGAGGCTGTCCAGGACGATCACCTCATCGGCGGGCACGTCGGGGGTACGC GCGCAGGTGGCCCCCGCTCCCAACCGCCCCAACGGCGCGCATCCCACTCCGACAGGTCCTGCTAGTGGAGTAGCCGCCCGTGCAGCCCCATGCC Ø, GAAGTGGACGATGGCGTCCACGCCGCGCAGTTCCCGGGCGAGGAGGCCGGCGTCGCGGATGTCGCCGTGGACGAAGCGCAGTCGCGGGTC CTTCACCTGCTACCGCAGGTGCGGCGCGTCAAGGGCCCCGCTCCTCCGGCCGCAGCGCCCTACAGCGGCACCTGCTTCGCGTCAGCGCCCAG GCACTGGAGCAGCGTCTGCGTGCCCTGCACGTTGGTCTCGGTGAACACGCGCGCCCCGCGATGGAGCGGTCCACGTGGCTCTCGGCCCCC 123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890 Д × Ø 团 œ Д ഗ Н 8 BsaAI Pm] I æ > H Д > [t4 ď Ω > Ω ĸ H 凶 ഗ O > Д Н Н ď > Ö R ы Ä Ω Ω ഗ Ø ഗ Ö > H Н Œ Ы E٦ Н ď ⋈ ഠ ď H œ G 囟 Z H Z > R Ø œ Ą O Ö > z H Alwni Ω Ы Ø ď K Н Д Ы > Ы Д 二 Ø Ø U [I4

sugar.finalgene b-1 Sequence

gggccgctcctccacgccgrccttcacgctccgcctactrctccgccgtcgagcccagtcgtcgtcttcgcc<u>ctg</u>acggcafgaaggccagc CCCGGCGAGGAGCTGCCGCACGAAGTGCGAGCCGATGAAGCCCGCACCTCCGGTCACCAGAAGCCGCACTGCCGTCTTCCTTTCGGTCGC Draili

ድ Ц Н > H Ö Ö Ø Ö [z, Н Ö ഗ 出 ſĽι > 24 Ø H

Ы

æ

ט

		_		
1	40	/1	64	

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	
MluI Aflili SfcI	!!
▼ GCTGTAGGTCGCGGTGTGGGTCGCACTGTCGGTGGCGGTGCGGGTCGCGGTGTGGGTTCGCACTGTCGGTGGCGCTGTCGGTGGCGTTCGTCGTGGGAAC CGACATCCAGCGCCACCCAGCGTGACAGCCACGCCAC	0606
BBABI GCGTCGCCGCGAGGTGCCCTCACCTCCCTCGCGGCCGCCGATCTCCATCAGATAGCTGCCGTACTCGGTGCGGGAGAGGCCTTCT CGCAGCCGCCGCCTCCACGGGAGGCCCCCCGGCCGCCGCCCTACTCCGGTGCCGCCCTCTCCGGAAGA CGCAGCCGCCCTCCACGGGAGGGAGCGCCGCCGCCCTACTCCACGCCCTCTCCCGGAAGA PAGBRGCCGCCAGGCCCCTCCACGGGAGAGAAGA CGCAGCCGCCCCTCCACGGGAGCCCCCTCCCCGCCGCCCCTACACACAC	9180
StuI CCCAGGCCGTGACAGGCGTCGATGAAGCCCATGCGGAAGGCGATCTCCTCAAGGCCCGCGATCCAGACGCCCTGCCGCTCCTCC GGGTCCGGCACTCCGGAGCGAAGGCGAAGGCGATCCGGGGCGCTAGGCCTAGGGCCCTAGGGCGCGAGGGGGGGG	9270
PflMI AlwNI AGACCTGGACGTACTGGGCGCCCGCAGGCGAGTCGTGGCTGTCCAGGCCAGGCGAAGCCGCGCGCCCAGGTTGACGAGTTCG TCCTGGACCTGCATGACCCGCCGGGCGTCCTCGCTCAGCTCACGCCACGGCCCAGGTTCAAGC TCCTGGACCTGCATGACCCGCCGCGCGTCTCAGCTCAACCCAACGCCAACTGCTCAAGC TCCTGGACCTGCATGACCCGCCGCGCGTCTCAGCTCAACCCAACTGCTCAAGC TCCTGGACCTGCATGACCCGCCGCGCGCTCAACTCAA	9360

141/164

12345678901234567890123456789012345678901234567890123456789012345678901234567890	
Aatii GCCCGGCCCCGCTCCAGGTAGACGCGGTGATCTCCCAGCTCGCCGCGGCGAGGGCCGGATGTTCTTGGCGATGTCGACG CGGGCCGGGGCGAGGTCATCTGCGCCACTGCACTAGAGGTCGACGCCGCCGGCCG	9450
AALII **ACGTCGTTGTCGTAGAGGCCGGTGACGCGAGGTTGGAGCGCGCTTGACGGGCTTCTCGACGAGGTCGGTC	9540
EarI ▼ GCGTCCACCTCGGCGACGCCGTACCGCTCGGGGTCCTTGACCGGGTAGCCGAAGAGCACGCCGCGCGCG	9630
ApaI ▼ AGGAGCGTGTAGAGGCCCGGGCAGATGTTGTCGCCCAGGATCAGGGCGCAGGTGTCGTCGCCGATGTGCTCGGCTCCGACGAGA TCCTCGCACATCTCCGGCCCGGGCACCTTCTACAACAGGGGGGTCTCCTAGTCCACAGCGGGCTACAGGCCGAGGCTGCTCT L L T Y L G P G H F I N D G L I L A C T D D G I H E A G V L	9720

142/	164
------	-----

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901	70 80 90 789012345678901234567890	
Tfil AGTGCGTCCGCGATTCCTGCGGGCTCTTTCTGGACCGCATAGTCGAGTTCTATTCCCAGGTGCCTGCC	EarI T CGTTTCCGAGAAGCGACTGGAAG 3GCAAAGGCTCTTCGCTGACTC G N G L L S Q F	9810
Tfil BsaBI NruI VY AGTTCGATGTGCTGGGGGGTCGAGATGATTTGAATCTCGCGAATACCGCCGAGCATGAGAACCGACAGCGGATAGTAGATCATCGGTTTG TCAAGCTACACGACCCCCCAGCTCTAAAACTTAAGAGCGCTTAATGGCGGCTCGTACCTTTGGCTGTCGCCTATCATCTAGTAGCCAAAC L E I H Q P I S I I Q I E R I G G L M L V S L P Y Y I M P K	AGCGGATAGTAGATCATCGGTTTG ICGCCTATCATCTAGTAGCCAAAC L P Y Y I M P K	0066
Tfil BstBI TGTAGACCGGAAGAATCTGCTTCGAAATGACCGAGGTCGCCGGCCG	CCGGCCAGGACTATTCCCTTCATT GGCCGGTCCTGATAAGGGAAGTAA G A L V I G K M	0666
MaeI BfaI BfaI ▼ CTCGGAAACTAGCAGCAGCGCCGCTGATAACGGTCGCCACGCCTCAATCCCCCCCC	FspI ▼ CTGCGCAGGGGAGTGTCACCACC SACGCGTCCCCTCACAGTGGTGG	10080

FIG. 32BB

ഷ

Ö

DLD

 $\boldsymbol{\wp}$

œ

R G

⊣

S

1	43,	/1	64	
---	-----	----	----	--

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	40 50 7890123456789012	50 789012	3456	60 7890123456	70	80	90		
BBLXI BBLXI ApaI ApaI CCTTTGGGGGGGGAAAAAAAAAAAAAAAAAAAAAAAAA	ApaI * * * * * * * * * * * * * * * * * *	GACGGCCGGG	Apai CCCTCAGGTG	GGGGATCGI)ටටටටටටට\	ATCGGGGGG	ATCGGG TAGCCC	10170	
GGTCAGCGCAGGA CCAGTCGCGTCCTT R L I	Pvuli ▼ GCGGGTGCGGGTCAGCGGGGCCTCCTCCCAGCCGTCCGCGGCGTCGCGTCCAGCTGGTTCAGGCGGGGGGGTGACGACC CGCCCACGCCCAGTCGCCTTCGGCGCCCCGGAGGGTCGGCAGCGCGCCGCCCCCCCC	TCCCAGCCGT AGGGTCGGCA E W G	Pvuli ▼ rcccagccgccgcgcccccagcTccagcggcggcggTgacgaC(AGGGTCGCCCGCCAGCTCCAGCTCCAGCGGCGGCGTGACGACGTCGGCCCGCCGCCAGCGCGAGGTCGACCAAGTCCGCCCGC	GCGCTCCAGG	Pvuli Cregricaec	CGGGCGGTGZ CGCCGCCACT RAT	ACGACC FGCTGG V V	10260	143/16
Scal GATCGAAGCCGTCCATGAAGTACTCGTCGCCC CTAGCTTCGGCAGGTACTTCATGAGCAGCGGC Q D F G D M F Y E D G	H K	ACGGCCGCCA TGCCGGCGGT V A A	ECONI CGACGGCCACCTCGCCGCCGCGCTCGACGAAGTCCCTGACGACCTCGGTGAGG GCTGCCGGCGGTGGCGGCGGCGCGGC	GCGCTCGACC CGCGAGCTGC R E V	PAAGTCCCTC CTTCAGGGA(F D R	SACGACCTCG	Econi GTGAGG CACTCC T L	10350	•
3GGGTCACGCGGCC	Afliii GAGGTGTCGGGGGTCACGCGGCCCCGCGATGTAGCGGGTCGCGCCGTCCAGGTCGGGGGAAGCCGGCCTCGCGGTACAGGTACACGTCGCCG CTCCACAGCCCCCCAGTGCCCCGGCGTACATCGCCCAGCGCCGGCCAGGTCCAGGCCCCTTCGGCCGGAGCCCCAGTGTCCATGTGCAAGCCGGCC	GTCGCGCCGT	CCAGGTCGGG	GAAGCCGGCC	TTCGCGGTA(BAGCGCCATC	Aflili * AGGTACACGTC	II FCGCCG AGCGGC	10440	

RFPI	ACE	MENT	SHEET
IVELL	ハしに	AITIAI	

Sequence
algene b-1 !
sugar.fin

		144/164		REPLACEN
	10530	10620	10710	10800
10 20 30 40 50 60 70 80 90 90 12345678901234567890123456789012345678901234567890	AGGAGATCGACCTGCACCGCGCTGCGGGTGGGCCGCATGGTGGCGGGCTTGATCCGCAGCAGTTCGGCGTCGGCCCCGGTG TCCTCTAGCTGGACGTGGCGCCCACGCGCCCACCGGCGTACCACCGCAACTAGCCGTCGTCAAGCCGCAGCCGGGGCCAC L L D V Q V A V Q P H A T P R M T A P K I R L L E A D A G T F8pI	CGCAGGCTGTTCAGGGCGTAGCCGTAGTGGAGTCCGGGGGTGCGCTCGCGGACCCGCTCCTCGAAGGCGTTGAGGGCCTCCTGG GCGTCCGACAAGTCCCGCATCGGCATCAGCTTACAGCCCCCCACGCGAGGGCGCTGGGCGAGGAGCTTCCGCAACTCCCGGAGGACC R L S N L A Y G Y D I H L G P T R E R V R E E F A N L A E Q	Sfei Nrul Prul Totale Control Prul Totale Cotale	CCGGCGTCGTTGAGGTCGGTCGACCCAGGTCGAGCCAGGCGGGGCCCCGGGGAGCCACGATGTAGGCGAAGCCGAGGTTG GGCCGCAGCAACTCCAGCTACTTCAGCTGGTCCAGCTCGTCCGTGCCGGGGCCCTCGTGCTACATCCGCTTCGGCTCCAAC G A D N L D A I F D V L D L L R S A R G P L V I Y A F G L N

145/164

Sequen
gene b-1
sugar.final

ဥ

Alwni PruII ATCGGCGACTCGCGCGCGCGCGCGCGCGCGCGCGGCGGCGGAAGGCGGCG
SCGCAGGTTCTCGCGGACGCGCGCGCCTCCAGGCCCTGCGCCTGCCCTGCGCCTGCCGCCTGCGCCTGCCGCC
AGCGG TCGCC F R T R

1	46	/	1 6	64	
---	----	---	-----	----	--

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	
NSPHI GCGGACGGGTCGTAGCGGCGCCCGGTCACACGGACGAGGCGGGAACATGCAGGTCGGGCCGGGGTAGAGGCCGACGCTGTAC CGCCTGCCCAGCATCGCCCGGCCAGTGTGCGTCTTCACCGTCGAGCCCGGCCCCAGTCTCCGGCTGCGACATG A S P D Y R A G T V R V C F H C R F M C T P G P Y L G V S Y	11250
BDSI W GGGAAGACGGGCTTCCTGGCGAGCGCGCGCGCGCGCGCGC	11340
Pfimi Primi Processora de Controlo Primi Processora de Controlo Primi Pr	11430
CGGGCCTCCTCCAGCGGGTGAAGGGGCTGTTGCCGTAGCGCACGGGGCGAACGAGGTGGCGGGGGGGG	11520

团

нроскурвц

ø

ᆸ

ρ, H H 闰 α Н ø K ĸ > П œ. ы Д ы Q H

147/164

sugar.finalgene b-1 Sequence

FIG. 32GG

Д

H

ď

æ

ĸ

н

œ

Ø Ы

二

z

Ö

ſĸ, Ω

ы

Д

O

Ö

വ

Д

4

Ö

Ö

H

Z

ď

×

Σ

ഗ

얺

团

12240

BamHI

TCATCTTGGCGTTGGTCCCGCCGGGGGCTGCCGCCGGGCAGGTCGAAGCCGAAGTTGTGGAGGGCGCGGGATCCGGGCGCGCGAGGTCCG

SERIAL NO.: 09/988,384 148/164 12150 11970 12060 CGTCGATCTCCGACGATCACGTACTGGTGGTTGTTGAGGCCGTGGCGGTCGTGGTCGCGCGACGACGACGCCGGGGAGGTCCGCGAGGTCCGCGAGGTCCT GCAGCTAGAGCTGCTAGTGCATGACCACCAACAACAACCGCCACCGCCAGCACCAGCCGCTGCTGCGGCCCCCTCCAGGCGCTCCACGA GGCAGCCCGGCGAGAAGTAGGCGCGCGGGTGTGCCACGCCTTTCAGGACCTCCATGACGAGGTCGCGGTGGATGCCGGTGGTGGTGGTGGT CCGTCGGGCCGCTCTTCATCCGCGCCCCACACGTGCGGAAGCCGGAAGTCCTGGAGGTACTGCTCCAGCGCCACCTACGGCCACCACCACCACCACCACCACCACCACCAC 回 Ø ㅁ ᆸ ď E ď ď H ď Ω Ö ď Ы н Sty1 Σ ρ, H Ö Ö œ Ω > Н ᆸ H H > > ഗ æ Ы Σ Д Ω 臼 Aflili MluI Ħ ø > Ω Ŀ Ы Д × œ ď Ή 团 O > 团 ы Н Ö Ω Z Apali > Ä z œ 二 Z E Q œ 跘 BsaAI Ø × ĸ > × Z H Œ ø, ល > ď 囟 ρι × O Н

Ω

团

O

H

123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890

Д

12600

149/164

12510

12330 12420 PvuII CGTCGTCGGTGACGACGCCCCCCCCCCCCCGAAGGCGTTGACGGCCTTGGTGGCGTGGAAGCTGAAGACCTCGGCGTCGCCGAGGCTGCCGG 123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890 GCAGCAGCCACTGCTGCCGCGGCGGGAGCTTCCGCAACTGCCGGAACCACCGCACCTTCGACTTCTGGAGCCGCAGCGGGTCCGACGGCC Ö S ы Ö Ω ď 团 > [Ł ល ſ۲, 二 Ø ₽ Styl × ď > Z ď ſι 囟

Ö

Ö

ď

>

>

[-Ω

Ω

ď

8

8

8

20

4

8

sugar.finalgene b-1 Sequence

CGGGCCGGCCGTCGACCGCGCAGCCGAGGGCGTGCGCGCGTCGAAGTACAGCCGCAGCCGTGCTCGTCGGCGACCTTCCGCAGCTGGT GCCCGGCCGGCAGCTGGCGCCTCCCGCACGCGCCGCAGCTTCATGTCGGCGTCCGGCACGAGCAGCCGCTGGAAGGCGTCGACCA Ø H œ × > ø Ω 闰 出 O Ы 卍 Ы × ſτι Д ď Þ 二 ď Ы Ö Ö K > Ω Ö 足 Д

ď

CGGCGCCCAGGCCCCCCAGAGGTGGACGCCGACGACGACGACGAGGTGCGGGGTGTGACCGCGGCGGCCACCTGGTCCGGGTCGAGGT <u> GCCGCCGCGTCCCCGCCGGGGTCTCCACCTGCGGCTGCTGCCGGCTCCACGCCCCAACTGGCGCCCCGGTGGACCAGGCCCAGCTTCCA</u>

Ы Ω ρι Ω Ø > 4 ď ø > ₽ Д ഷ Н ß ø > > Ö > 工 Н 3 Ö ĸ Д U ď ø

Ω

BbsI

TECCEGIGICCEGEICCAICICCECCAAGACCCGGGGIGAGCCCGAICCAGCGCAGIGCGTGCGGGGIGGCGAACGICAICGACGCA

ഗ Σ ₽ ſτι 4 ď E Д 二 ď R 3 H Ö Н H Q, > Œ K Ω Н Ω Д Д ₽ O Z

REPL	ACEN	MENT	SHEET
------	------	------	-------

sugar.finalgene b-1 Sequence	
finalgene b	edneuce
finalgene b	i.
g.	۵
g.	gene
sugar	g.
	sugar

	Ji	150/16	4	REPLACEMEN
	12690	12780	12870	12960
10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	Beli Bali Naphi TGATCACTTCGCCGGTGAGGCGGCGTGCCGGAGGAGCTGGAGCCGGCCG	Alwni Apai CGAGCCCGGCGACGCTCCTCGAACTCGCGGACGAGCGGCCGCTTGGACAGCCCACTGGCTGTCGAGGCCCGGTCGAGCCGCTCGT GCTCGGGCCGCTGCGGAGGAGGCGCCTGCTCGCCGGCGCGCGC	Bemi ACAGCCTGGCGCGGTCGATGCGGTTGGGCCGCCCCACGAGGAGCGGCTGGTCGAAAAGCGGCGGGCCGCCGAAGAATGCGAGGTCGATA TGTCGGACCGCCGCCAGCTACGCCGGGGGGGGGG	Tfil XmnI XmnI AGGCGCTTTTCACGGATGTTCCCTCCGGGCCACCGTCACGAAATGATTCGCCGATCCCGGAATCGAACGAGGTCGCCGCGCTCCACCG TCCGCGAAAAGIGCTTACTAAGGGCTTACTTACTAAGGGCTTAGGGCTTGCTCCAGCGGCGCGCGAGGTGGC L A S K V

151/164

REPLACEMENT SHEET

Þ 饵

Ø Ö

Ö Ц

Ω ы

Д Д တ

Ö

Ø

卍

10 20 30 30 30 30 30 30 30 30 30 30 30 30 30	
TGACGTACGACGAGATGGTCGATTGTGGTGGTCGATTTTCGGGGGACTCTAATCCGCGCGGGAACGGGGACCGACAAGAGCACGCTATGCGC ACTGCATGCTGCTCTACCAGCTAACACCACCAGCTAAAGCCCCCCTGAGATTAGGCGGCGCCTTGCCCTGGCTGTTCTCGTGCGATACGCG	13050
Bamhi Tfii TCTCGATGTGCTTCGGATCACATCCGCCTCCGGGGTATTCCATCGGCGCCCGAATGTGATGTCTTGACAGGATCCGGGAATCAGCCG AGAGCTACACGAAGCCTAGTGTAGGGGGCCCCATAAGGTAGCCGCCGGGCTTACACTAGGAACTGTCCTAGGCCCTTAGTCGGC	13140
BSAAI EALI Aflili ACCGCCGGGAGGGCCGCGCCGCTCCGCGGAAGAGTACGTGTGAGAAGTCCCGTTCCTCTTCCGGTTCCGTTCCGCTTCCGGCCCGG TCGGCGCCCCCCGCCGCGCGCCCTTCTCATGCACACTCTTCAGGGCAAAGGCCAAAGGCCAAGGCCAAGGCCGAAGGCCGAAGGCCGAAGGCCGAAGGCCGAAGGCCAAAGGCCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCCAAAGGCCCAAAGGCCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCAAAGGCCCAAAGGCCCAAAGGCCCAAAGGCCCAAAGGCCCAAAGGCCCAAAGGCCCAAAGGCCCAAAGGCCAAAAGGCCAAAAGGCCAAAAGCCCAAAAGCCCAAAAAA	13230
ECONI Apai	
TCTGGAGTTCTCCGTGCGCCGTACCCAGCAGGGAACGACCGCTTCTCCCCCCGGTACTCGACCTCGGGGCCCTGGGGCAGGATTTCGCGGC	13320

sugar.finalgene b-1 Sequence

AGACCTCAAGAGGCACGCGCATGGGTCGTCCCTTGCTGGCGAAGAGGGGCCCATGAGCTGGAGCCCCGGGACCCCGGTCCTAAAGCGCCCG

152/164

10 20 30 40 50 60 70 80 90 90 123456789012345678901234567890123456789012345678901234567890	20 7890123456	30 7890123456	40 7890123456	50 5789012349	60 67890123456	70 7890123456	80 789012345	90		
BsaBI				Dre	FspI DraIII					
CGATCCGTATCCGACGTACGCGAGGCTGCGAGGGTCCGGCCCACCGGGGGCACCCCCCGAGGGGGGACGAGGTGTGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	STACGCGAGAC LATGCGCTCTG YAR	TGCGTGCCGA ACGCACGGCT RAB	GGTCGGCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG	CACCGGGTC GGTGGCCCAC H R V·	CAGGGTCCGGCTGCGCACCCCCGAGGGGGACGAGGTGTGGCTGGTCGTCGTCGTCGTCGTCGTCCCCCAGGCTCCCCCTGGGGGCTCCCCTGCTCCACACCGACCAGCTCCCCAGGCTCCCAGGCTCCCAGGCTCCCAGGCTCCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTGCTAGAGGGGGGGACCAAGGCTGGCT	GGGGGACGAG CCCCCTGCTC G D E	GTGTGGCTG CACACCGAC V W L	GTCGT CCAGCA V V	13410	
CGGCTACGACCGGGGGGGGGGTCCTCGCCGATCCCCGGTTCAGCAAGACTGGCGCAACTCCACGACTCCCCTGACCGAAGCCGAAGCC GCCGATGCTGGCCCGCGCCCGCCAGGAGCGGCTAGGGGGCCAAGTCGTTCTGACCGCGTTGAGGTGCTGAGGGGGACTGGCTTCGGCTTCGG	SCGCGCCAGG	TCGCCGATCC	CCGGTTCAGG	AAGACTGG(STTCTGACCC	GCAACTCCACC	ACTCCCCTGA TGAGGGGACT	CCGAAGCCG	AAGCC	13500	•
	IHdsN				PÉIMI PvuII	L				
GCGCTCAACCACACATGCTGAGTTCCGAACCCGCGCGCGC	▼ ATGCTGAGTTC	CGAACCCGCC	GCGGCACAC	CCGCTGCG	CAGCTGGTGG	CCGTGAGTTC	ACCATGCGC TGGTACGCC	CGGTG	13590	

CGAGTTGCTGCCGCCCCGGGTCC GCTCAACGACGCCGGGGCCCCAGG